1-1-2012

The steroid receptor coactivator, SRC-1, can function as a transcriptional corepressor to selectively suppress anti-tumourigenic target genes in Tamoxifen resistant breast cancer

Claire A. Walsh
Royal College of Surgeons in Ireland

The steroid receptor coactivator, SRC-1, can function as a transcriptional corepressor to selectively suppress anti-tumourigenic target genes in Tamoxifen resistant breast cancer

Claire A. Walsh BSc (Hons), MPhil

A thesis presented to the Royal College of Surgeons in Ireland
St. Stephen's Green, Dublin 2

Submitted for Doctor of Philosophy

Endocrine Oncology Research Group
Royal College of Surgeons in Ireland

Supervisor: Professor Leonie Young
Co-Supervisor: Professor Arnold Hill

September 2012

Declaration

I declare that this thesis which I submit to RCSI for examination in consideration of the award of a Doctor of Philosophy is my own personal effort. Where any of the content presented is the result of input or data from a related collaborative research programme, this is duly acknowledged in the text such that it is possible to ascertain how much of the work is my own. I have not already obtained a degree in RCSI or elsewhere on the basis of this work. Furthermore I took reasonable care to ensure that the work is original and to the best of my knowledge, does not breach copyright law and has not been taken from other sources except where such work has been cited and acknowledged within the text.

Signed _____

Student Number _____

Date _____

Acknowledgements

As I began my PhD in RCSI four years ago, my cousin Teresa had just completed hers. Her parting words of wisdom to me were "A PhD will break you down but it will also build you back up". I can safely say that there have been many people who have helped to build me back up over the last four years and I owe them all a huge gratitude.

First and foremost to my supervisor, Prof. Leonie Young. Her door was always open for advice, support and scientific inspiration. I offer my many thanks to her for guiding me so well through my PhD experience. I also offer sincere thanks to my co-supervisor Prof Arnold Hill whose dedication to research is a constant source of strength to the group.

To the wonderful and talented post docs of the group, Marie McIlroy, Jean McBryan, Christopher Byrne and Aisling Redmond. I have learnt so much working alongside each of them and will be forever grateful for their constant help and support over the last three years. I offer special thanks to Marie and Jean for taking time out of their experiments to review my thesis, it was very much appreciated.

To Sinead Cocchiglia, who must have answered approximately one million questions for me over the years without ever getting annoyed. Her guidance has been invaluable and an inspiration. I also send my many thanks across the desk to Fiona Bane, the epicentre of the lab and a "wild good" scientist.

To Eamon, Damir, Sarah, Jarlath, Jane, Gemma and Katie, I am so lucky to have been able to do my PhD alongside such great people. It was a genuine pleasure to come to work every day and I will miss them all terribly.

To Dr. Damian McCartan, whose excellent research paved the way for this work. I offer my genuine thanks for all his help over the years. I also offer special thanks to Jarlath Bolger for all of his efforts and support on this project, it was very much appreciated. Also, thank you to Aoife Ni Chathalain, whose western blotting skills have proven invaluable. To Yuan Hao, who was always on hand with her wealth of bioinformatic knowledge to help me make sense of my data. It was a pleasure to work with her over the last year. To Paul Tibbetts, for his work on the TMA and for his statistical advice. I also extend my warmest thanks to Dr.

Sudipto Das, who took the time to teach me about the MeDIP experiment and who was a constant source of advice for all my methylation queries.

I would like to especially thank Prof. Jianming Xu and his research group at Baylor College of Medicine for welcoming me so warmly into their group and for being so generous with both their time and their data. I would like to particularly thank Dr. Li Qin for all her assistance.

I would also like to offer my warmest and most sincere thank you to the patients of Beaumont hospital who, during what is an undoubtedly difficult time, have been so consistently generous in their support of our research.

I would like to acknowledge and thank the postgraduate school of RCSI, Prof Kevin Nolan, Dr. Helen McVeigh, Bernadette Kearney and Cliona Lyes for all their efforts over the last four years. It was a pleasure to be part of such a well run programme. Also special thanks to Kay McKeon and Joan Ni Gabhann of MCT for all their help with the flow cytometer.

To Claire Wynne and Yvonne Smith, a girl couldn't ask for two better friends. I am so lucky to have had the two of them through all of this. To Cormac and all the other RCSI postgraduate students, past and present, it's been a brilliant four years with many, many fond memories.

To my wonderful family, at last at last I am finished college. Under no circumstances could I have made it this far without their love and support. I never had to walk one step of this journey alone, I love them with all my heart.

Finally to Frainc, who has held my hand through all the ups and downs of the past three years, and who always had a cup of tea and chocolate on standby for me after a long day in the lab. I can never thank him enough X

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

List of Tables

Summary

Introduction

Biological targeted therapies have become a mainstay of breast cancer treatment for many patients. Tamoxifen is one of the most successful therapies to date due its anti-estrogenic effect in the breast. However, a significant number of patients, almost 40%, develop resistance to Tamoxifen therapy and are thus non responsive to its effects or in some cases, the tumour can actually progress on Tamoxifen. The steroid receptor coactivator 1 (SRC-1) has been associated with the development of Tamoxifen resistance in cellular models and in clinical patient populations. Furthermore, Tamoxifen treatment can upregulate SRC-1 activity in a resistant environment. SRC-1 has been shown to form a non steroidal transcriptional complex with the developmental protein, HOXC11; together they work to assist the deregulation of anti-tumourigenic signalling pathways within a tumour cell. SRC-1/HOXC11 activity in endocrine resistance has been reported as a significant marker of poor prognosis in a clinical patient population. Currently, the secreted protein, S100β is the only identified target gene of SRC-1/HOXC11 activity in endocrine resistant breast cancer. Given the clinical relevance of this functional interaction, it is important to identify additional target genes in order to gain a greater insight into the effector pathways being regulated, as a tumour cell is driven from a responsive to a resistant state.

Hypothesis

SRC-1 has been upheld as a master regulator in cancer. It is well known for its ability to bind across unrelated families of transcription factors and coactivate the regulation of multiple genes in multiple complex physiological states. On the basis of SRC-1 ChIPsequencing (ChIPseq) and SRC-1 knockout DNA microarray data conducted in an endocrine resistant cellular model, this work hypothesises that SRC-1 can juxtapose its functionality between that of a coactivator and that of a corepressor depending on the transcriptional milieu. In this way, SRC-1 can have as significant an impact on tumour progression through the downregulation of "protective" genes as it could through the upregulation of its better characterised "progressive" genes. It is also hypothesised that HOXC11 is a common transcription factor of choice for SRC-1 in both transcriptional scenarios.

Results

Global analysis of SRC-1 in conjunction with a SRC-1 knockout DNA microarray identified 1,061 genes which were directly suppressed by SRC-1 activity in Tamoxifen driven endocrine resistant breast cancer. The luminal marker, CD24 was identified as a significant target gene of SRC-1 negative regulation. The developmental protein, HOXC11, is recruited with SRC-1 to

negatively regulate CD24. Additional target genes of SRC-1/HOXC11 activity were subsequently identified via inclusion of HOXC11 ChIPseq data. The tumour suppressor gene, PAWR, was also shown to be negatively regulated by SRC-1/HOXC11 activity in the same resistant context as CD24. These molecular observations were confirmed in a clinical patient population where loss of CD24 and PAWR associated with poor disease free survival.

Conclusion

Both CD24 and PAWR expression associate with the well differentiated, tightly controlled processes of cell adhesion and apoptosis, respectively. Targeted downregulation of proteins such as these, deregulates a cell and makes it increasingly vulnerable to tumourigenic signalling mechanisms. Together, SRC-1/HOXC11 decommission important anti-tumourigenic pathways to induce a deregulated cellular state which is ultimately more vulnerable to the intentional and targeted pressure of uncontrolled growth signalling pathways.

Publications and Communications

Publications

Walsh CA, Qin L, Tien JCY, Young LS, Xu J. The Function of Steroid Receptor Coactivator-1 in Normal Tissues and Cancer. Int J Biol Sci 2012; 8(4):470-485.

International Communications

April 2012 - The role of the steroid receptor coactivator SRC1 and its functional partner HOXC11 in the development of endocrine resistant breast cancer.
International Association of Breast Cancer Research, Manchester, UK. **(Poster presentation)**

November 2011 - The role of the steroid receptor coactivator SRC1 and its functional partner HOXC11 in the development of endocrine resistant breast cancer
San Antonio Breast Cancer Symposium (SABCS),Texas, USA. **(Poster presentation)**

September 2009 - HOXC11: Establishing communication between steroid and developmental transcription factor pathways in breast cancer progression
EMBO Molecular Medicine Workshop, Turin, Italy. **(Poster presentation)**

National Communications

March 2012 - The role of SRC-1 and its functional partner HOXC11 in the development of endocrine resistance in breast cancer
Irish Association of Cancer Research (IACR) Conference, Belfast, Ireland. **(15 minute oral presentation)**

March 2011 - The role of SRC-1 and HOXC11 in tumour adaptability
Irish Association of Cancer Research (IACR) Conference, Cork, Ireland. **(Oral poster presentation)**

Abbreviations

4-OHT	4-hydroxytamoxifen
AD	Activation domain
AF	Activation function
AI	Aromatase inhibitor
ATCC	American Tissue Type Collection
bp	Base pair
Bcl-2	B-cell CLL/lymphoma 2
BSA	Bovine serum albumin
ChIP	Chromatin immunoprecipitation
ChIPseq	Chromatin immunoprecipitation sequencing
CK	Cytokeratin
DAXX	Death domain associated protein
DBD	DNA binding domain
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
ECM	Extracellular matrix
ECO	European Cancer Observatory
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EMT	Epithelial mesenchymal transition
ER	Estrogen receptor
ERE	Estrogen response element
ES	Embryonic stem cell
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
FCS	Fetal calf serum
FSC	Forward scatter
FDR	False discovery rate
FGF	Fibroblast growth factor
FISH	Fluorescent in situ hybridisation
GFR	Growth factor receptor
GPI	Glycosylphosphatidylinositol
GR	Glucocorticoid receptor

HAT	Histone acetyltransferase activity
HDAC	Histone deacetylase
HMT	Histone methyltransferase
HER	Human epidermal growth factor receptor
Ig	Immunoglobulin
IGF	Insulin growth factor
IHC	Immunohistochemistry
LB	Lysis buffer
LBD	Ligand binding domain
MAPK	Mitogen activated protein kinase
MEM	Minimum essentials medium
METABRIC	Molecular Taxonomy of Breast Cancer International Consortium
mL	Millilitre
MMTV-PyMT	Mouse mammary tumour virus-polyoma middle T
mRNA	messenger RNA
n	Experimental number
NCRI	National Cancer Registry of Ireland
NLS	Nuclear localisation signal
NR	Nuclear receptor
NRID	Nuclear receptor interacting domain
p	p-value
PBS	Phosphatase buffered saline
PI	Protease inhibitor
PI3-K	Phosphoinositide 3-kinase
PIN	Prostatic intraepithelial neoplasia
PITX2	Pituitary homeobox transcription factor 2
PKC	Protein kinase C
PMT	Photo multiplier tube
PR	Progesterone receptor
PTEN	Phosphatase and tensin homolog
RISC	RNA induced silencing complex
RNA	Ribonucleic acid
RT	Reverse transcription
RTK	Receptor tyrosine kinase

SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
siRNA	small interfering RNA
SERD	Selective ER downregulators
SERM	Selective ER modulators
SRC	Steroid receptor coactivator
SSC	Side scatter
TBS	Tris buffered sulphate
TF	Transcription factor
THAP1	THAP domain containing, apoptosis associated protein 1.
TMA	Tissue microarray
TR	Thyroid receptor
TSS	Transcriptional start site
TTS	Transcriptional termination site
v/v	Volume by volume
VEGF	Vascular endothelial growth factor
w/v	Weight by volume
WHO	World Health Organisation
WT-1	Wilms' tumour 1
µL	Microlitre
µM	Micrometer
ZIPK	Zipper interacting protein kinase

1. General Introduction

1.1. Breast cancer

Breast cancer is the most frequent cancer among women with an estimated 1.38 million new cancer cases diagnosed each year (23% of all cancers), and is ranked second in terms of all cancers at 10.9%. As of 2008, incidence rates of breast cancer are high in the developed regions of the world (except Japan) and low in most of the developing regions [1]. Breast cancer is now recognised as a worldwide public health issue and has been the focal point of immense research efforts and clinical advancements within the last number of years. As a result of this, there is now a varied repertoire of surgical procedures and sophisticated therapies to tackle a breast cancer diagnosis, so that each patient's treatment can be tailored to the type, stage and extent of their cancer. Despite these important advancements, breast cancer is still the most frequent cause of cancer death in women in both developed and developing regions [1].

1.1.1. Breast cancer in Ireland

The most recent report from the National Cancer Registry of Ireland (NCRI) was issued in May 2012 and reviewed incidence, mortality, treatment and survival of breast cancer in Ireland from 1994 -2009. The report stated that breast cancer accounted for 33% of all invasive cancers (excluding non-melanoma skin cancer) in Irish women, thus making it the most common tumour diagnosis for Irish women. It also concluded that breast cancer was the second leading cause of cancer death in Irish women, after lung cancer at 16% of the total female cancer deaths [2]. Comparisons made by the European Cancer Observatory (ECO) for 2008, show that Ireland has a high incidence rate of breast cancer and has the 4th highest breast cancer mortality rate of 30 European countries (Figure 1.1) [3].

Retrospective comparisons from the 2012 NCRI report, that Ireland has been less successful in tackling breast cancer in comparison to other European countries. However, significant improvements have been reported on a national scale with regards to the diagnosis and treatment of breast cancer in Ireland. One of the most important events to impact on these improved breast cancer statistics for Ireland, was the introduction of a national screening programme. BreastCheck is a government funded initiative which was set up in the year 2000 and offers free mammogram screening for women who are between the ages of 50 and 64 as half of the Irish women diagnosed with breast cancer were aged between 45 and 64 years. The advent of BreastCheck in Ireland coincides with an increase in the incidence rate of breast cancer in Ireland since 1994, specifically in the age bracket of 50-64 years. Importantly though, the NCRI report also presents evidence of a steady decline in breast

cancer mortality and a significant trend towards increased survival over the same time period, which attests to the benefits of a national screening programme for women who are at a higher risk of breast cancer.

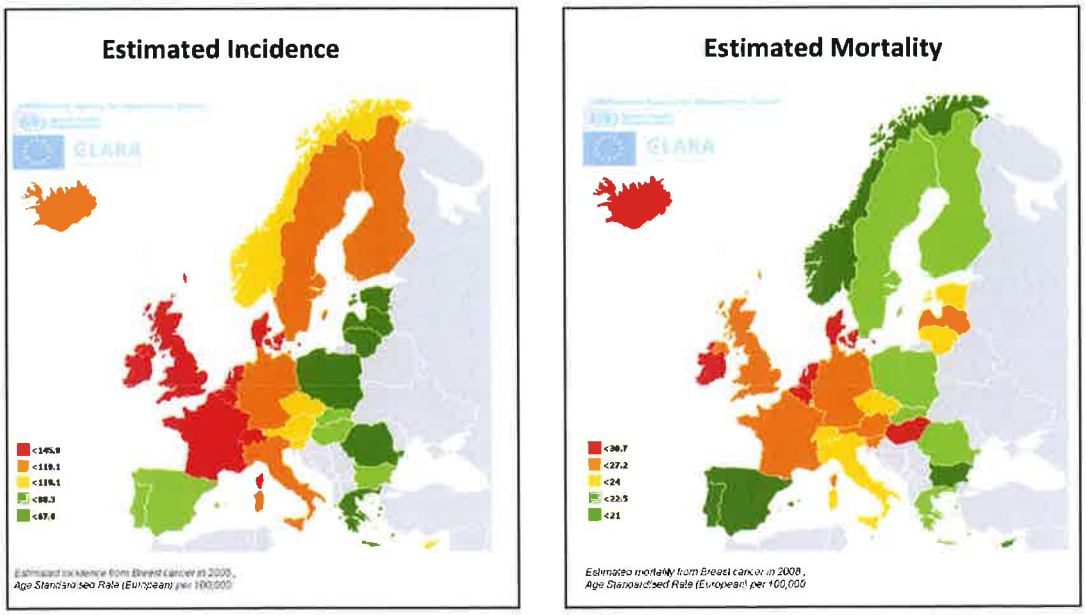


Figure 1.1. Estimated incidence and mortality of breast cancer within the EU in 2008.
Adapted from Ferlay et al, 2010

1.1.2. Epidemiology of breast cancer

Breast cancer is in essence a heterogenous disease which associates with various clinical outcomes. To determine whether the aetiology of breast cancer was due to genetic or ethnic factors, migrant studies were undertaken to record the incidence of cancer among emigrants/immigrants. The study determined that the incidence of breast cancer varied with the changing environment thus underpinning the relevance of environmental factors such as life style and socioeconomic conditions [4]. A number of these environmental risk factors have been well described; reproductive behaviour, age at menarche and menopause, estrogen intake, body mass index (post menopausal), alcohol consumption and exposure to ionising radiation, all have associations with an increased risk of breast cancer [5]. There is also a genetic element to breast cancer risk where strong associations have been made between age as well as a prior history of familial breast cancer. Specific germ line mutations in the BRCA1 and BRCA2 genes have also been identified as a serious genetic risk factor for breast cancer as well as ovarian cancer. Carriers of these mutations, whilst rare, carry a 40% to 85% lifetime risk of developing breast cancer [6]. High penetrance genes

account for only 25% of the familial risk of breast cancer. Therefore it is anticipated that the remaining breast cancer familial susceptibility is due to a relatively large number of low penetrance genes [7].

1.1.3. Breast cancer subtypes

A seminal study based on the expression patterns of over 500 genes identified four main molecular classes of breast cancer; luminal A, luminal B, HER-2 overexpressing and basal-like [8, 9]. Further independent studies have confirmed the authenticity of these intrinsic subtypes in both human primary cultures and in cell lines therefore the nomenclature attributed to them has become commonplace in breast cancer associated research [9-12]. These intrinsic subtypes accurately reflect the weighted role of epithelial luminal cells in breast carcinogenesis as approximately two-thirds of diagnosed breast cancers were classified as either luminal A or luminal B subtypes [8].

Microarray studies have shown that luminal types of tumours express high amounts of luminal cytokeratins and genetic markers of luminal epithelial cells of normal breast tissue (Figure 1.2) [13]. Luminal A and luminal B subtypes share many of the same clinical parameters; they are both positive for estrogen (ER) and/or progesterone receptor (PR) and both express many of the same ER-associated genes. However, luminal A subtypes often present as low grade and are considered to be well differentiated tumours as they often express high levels of both ER and PR. Since luminal subtypes are predominantly hormone receptor positive, they are more often than not given endocrine targeted therapies such as Tamoxifen or in the case of post menopausal women; aromatase inhibitors (AIs). In particular, luminal A subtypes respond well to such therapies and display relatively favourable survival times [14]. In contrast, luminal B cancers express lower levels of steroid hormone receptors and may also express growth factors such as the human epidermal growth factor receptor (HER2). Luminal B subtypes can also be characterised by high expression levels of proliferative genes such as CCNB1, MKI67, and MYBL2 [15]. Overall, the luminal B cancers represent a clinically distinct group with a different and worsened disease course than their luminal A counterparts, in particular with respect to relapse. It has been suggested that the luminal B subtype reflects a group of patients who will not benefit from adjuvant Tamoxifen despite positive receptor values [16]. The potential clinical significance of this molecular subtype is further highlighted by the similarities in expression of some of the luminal B genes with the ER-negative tumours in the basal-like and HER2/ERBB2+ overexpressing subtypes, thus suggesting that high levels of these overlapping genes is associated with poor disease outcome [17].

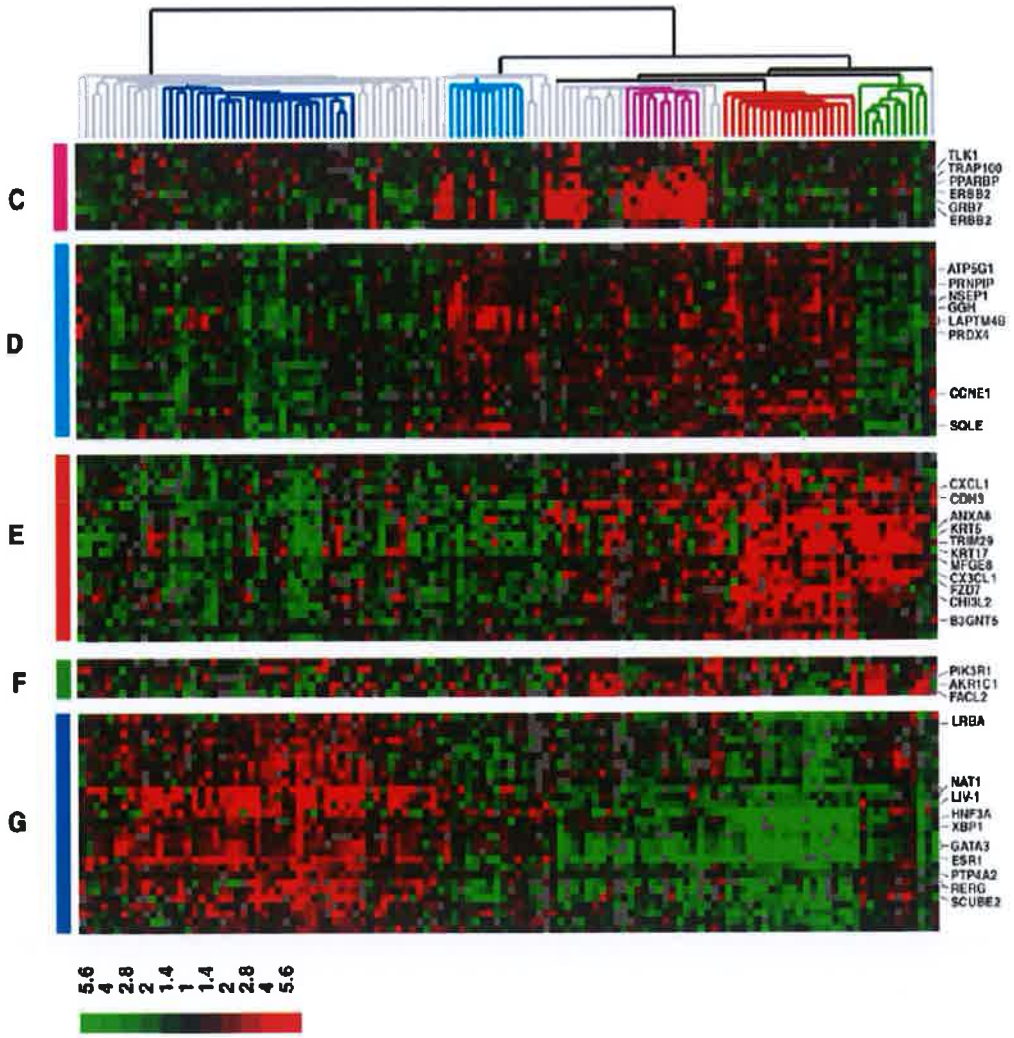


Figure 1.2. Molecular classifications of intrinsic breast cancer subtypes. Hierarchical clustering of 115 tumour tissues and seven non-malignant tissues using the "intrinsic" gene-set. Experimental dendrogram showing the clustering of the tumours into five subgroups. Branches corresponding to tumours with low correlation to any subtype are shown in grey. (C) Gene cluster showing the ERBB2 oncogene and other co-expressed genes. (D) Gene cluster associated with luminal subtype B. (E) Gene cluster associated with the basal subtype. (F) A gene cluster relevant for the normal breast-like group. (G) Cluster of genes including the estrogen receptor (ER) highly expressed in luminal subtype A tumours. Scale bar represents fold-change for any given gene relative to the median level of expression across all samples. *Adapted from Sorlie, 2004 [16]*

HER2 or ERRB2-overexpressing cancers are the third major breast cancer subtype. Essentially these tumours overexpress the HER2/ERBB2 growth factor receptor and the ERBB2-associated genes. They are however hormone receptor negative and do not express genes that define the luminal subtype [15]. HER2-overexpressing cancers are highly

aggressive in nature however they can respond well to targeted therapies such as Trastuzumab (Herceptin™) or Lapatinib.

The basal subtype represents approximately 15% of breast cancers, this subtype is commonly characterised as a proliferative, high-grade tumour with a poor prognosis [8]. Specifically, basal-like breast cancers do not express ER, PR, or HER2. These cancers lack expression of luminal cytokeratins, smooth-muscle–specific markers, and certain integrins which usually associate with myoepithelial cell populations. In contrast, some of these basal-like cancers, exhibit high expression of "basal" cytokeratins such as CK5 and a variety of growth factor receptors, including high levels of epidermal growth factor receptor (EGFR), c-kit (a tyrosine kinase in breast epithelium), and growth factors such as hepatocyte growth factor and insulin growth factor (IGF) [9, 18]. Basal-like tumours have also been associated with dysfunction of the BRCA1 pathway caused by *BRCA1* gene promoter methylation or BRCA1 transcriptional inactivation, or both [19, 20].There is no specific therapy currently targeted at basal tumours. Chemotherapy represents the sole available systemic treatment, and although basal tumours are relatively chemosensitive, its benefit remains limited [14].

In 2012, METABRIC (Molecular Taxonomy of Breast Cancer International Consortium) conducted the largest breast cancer study to date and reclassified breast cancer into 10 molecular subtypes based on an integrated analysis of copy number and gene expression (Figure 1.3) [21]. This landmark study assembled a collection of over 2,000 clinically annotated primary breast cancer tumours. 997 of the samples were analysed as a discovery group and a further 995 were used to validate the reproducibility of the discovery data. The 10 integrative clusters (IntClust) were typified by well defined copy number aberrations and split many of the original intrinsic subtypes. Two of the subtypes associated with a favourable outcome in breast cancer were IntClust3 and IntClust4. The IntClust3 subtype has very few copy number aberrations and had quite low genomic instability. These tumours were predominantly luminal A cases and enriched for histotypes that have a typically good prognosis. The IntClust4 encompasses ER-positive and ER-negative cases and varied intrinsic subtypes, this subgroup has very few aberrations but does exhibit extensive lymphocytic infiltration. IntClust7 and IntClust8 were also associated with favourable outcome and a luminal A subtype. IntClust1 has a predominantly ER positive cancer population however, it is classed as a luminal B subgroup with an intermediate prognosis. IntClust2 represents a high risk ER-positive luminal subgroup in which almost all tumours had a high frequency of amplification for cyclin D1 (CCND1). The majority of basal like tumours formed a stable,

mostly high genomic instability subgroup referred to as IntClust10. Finally, IntClust5 accounts for the ERBB2-amplified cancers and is composed of HER2 enriched ER negative and also includes HER2 enriched ER positive cases [21].

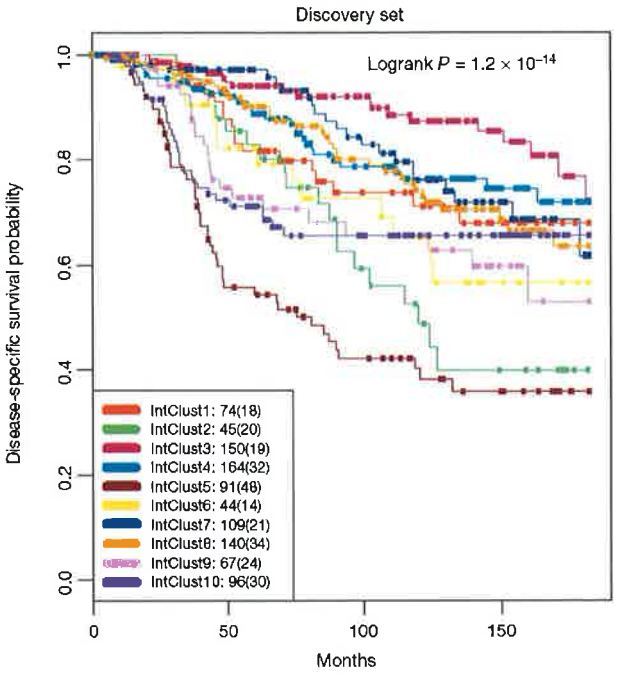


Figure 1.3. The 10 novel breast cancer subtypes with distinct clinical outcomes. Kaplan–Meier plot of disease-specific survival (truncated at 15 years) for the 10 novel integrative subgroups identified in the METABRIC cohort. For each cluster, the number of samples at risk is indicated as well as the total number of deaths (in parentheses). *Adapted from Curtis et al, 2012 [21].*

Whilst the original intrinsic subtypes from Perou et al, are prominent in the clinical setting, this new study emphasises how breast cancer is simply an umbrella term for numerous subgroups of disease and how important accurate treatment regimes and personalised targeted therapies will be to improving patient prognosis in the future.

1.2. Endocrine signalling in the normal breast

Normal cellular function is meticulously controlled by a series of biological networks. These networks are highly complex with multiple levels of regulation, fine tuning capabilities, redundancy, and evolvability, all of which are necessary to keep the cell alive under conditions of stress, toxicity or hostile environmental factors. Cancer cells exploit these normal functions, which are often altered genetically during oncogenesis, to provide them with a survival advantage and the ability to escape the effects of treatment. The ER

signalling pathway is an example of a complex biological pathway that controls a variety of functions, such as cell proliferation, apoptosis, invasion, and angiogenesis, and is exploited by breast cancer cells to serve as a major survival pathway driven by the female hormone estrogen [22].

Since almost 75% of breast cancers express ER and are dependent on estrogen for survival, an understanding of the normal signalling mechanisms of the ER pathways and its associated target genes is essential.

1.2.1. Normal cellular structure of the human mammary gland

The mammary gland is unique from other organs with regards to its development. Unlike most mammalian organs which develop primarily embryonically with a more or less linear progression toward functional maturity, development of the mammary gland is primarily post-pubertal and progresses initially in a linear fashion but then undertakes a cyclical pattern of development once sexual maturity is reached [23]. The epithelium of the mammary gland exists in a highly dynamic state and undergoes dramatic morphogenetic changes during puberty, pregnancy, lactation, and regression. All of these changes are prompted and influenced by endocrine activity within the breast. This innate plasticity coupled with its heightened sensitivity to endocrine influence makes the mammary gland a vulnerable target for tumourigenic events.

The cellular structure of the normal mammary gland develops as a branching tree-like network of ducts lined by a double layer of epithelial cells that is surrounded by fibroblasts and embedded in an extracellular stromal matrix (Figure 1.4) [24]. The primary ducts which are located near the epidermal surface of the nipple are encased in myoepithelial cells on the exterior and lined with ductal luminal cells along its length. The ducts grow, divide, and form club-shaped terminal end buds. These terminal end buds give rise to new branches and small ductules known as alveolar buds. These alveolar buds consist of secretory epithelial cells that can undergo functional differentiation in pregnancy for milk production [25].

The cellular compartments within the mammary gland are not only distinguished by their relative positions, but by the proteins that they express. The myoepithelium expresses a distinct subset of epithelial cytokeratins (CK 5/6, 17 and 14), the filament protein; vimentin, the common acute lymphoblastic leukaemia antigen (CALLA) and smooth muscle actin [26, 27]. In contrast, the luminal cell type can be distinguished by expression of a subset of epithelial cytokeratins (CK 8, 18 and 19), ER, PR, the cell adhesion molecule; E-cadherin and low (but detectable) levels of milk proteins [26-28].The luminal cells also account for more

than 90% of epithelial cell proliferation that is observed in the non-pregnant gland [29]. Significantly, many of these proliferating cells express ER and PR indicating that the luminal cell type is a major target for breast tumourigenesis [30].

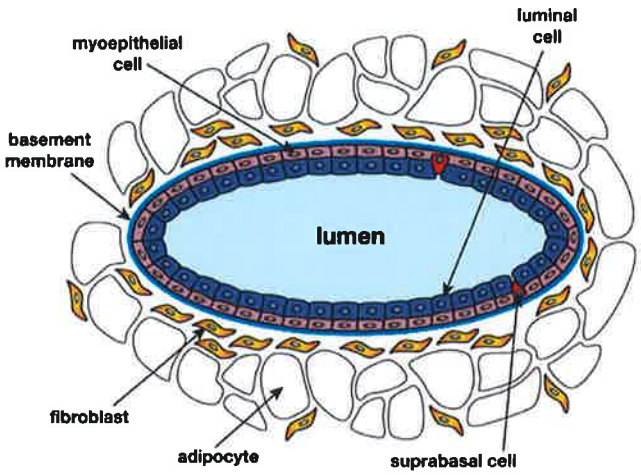


Figure 1.4. A schematic diagram showing the cellular structure of a functional mammary gland duct. The typical mammary gland duct consists of an inner luminal epithelial cell layer and an outer myoepithelial cell layer. The ductal lumen is formed as central body cells apoptose and outer cells differentiate into luminal epithelial cells. The ductal structure is separated from the surrounding extracellular stromal matrix by the basement membrane *Adapted from Visvader, 2009 [24].*

1.2.2. Estrogen

Estrogens refer to the naturally occurring steroid hormones which were named for their importance in the estrus cycle of humans. There are three major naturally occurring estrogens in women; estrone (E1), estradiol (E2), and estriol (E3). Estradiol is the predominant estrogen during the reproductive years both in terms of absolute serum levels as well as estrogenic activity[31]. During menopause, estrone becomes the most prominent circulating estrogen whilst during pregnancy it is estriol. Estrogens are produced primarily by the ovaries, and during pregnancy, they can also be produced by the placenta. Some estrogens are also produced in smaller amounts in the liver, the adrenal glands, the breasts and in the adipose tissue [31]. During menopause, estrogen production in the ovaries ceases and these secondary sources of estrogens become especially important. In postmenopausal women, estrogens are derived by aromatase conversion of testicular, adrenal or ovarian androgens [32].

1.2.3. The Estrogen Receptor (ER); Structure and Isoforms

Although estrogens were discovered in the 1920s, the ER was not identified until the 1960s. Estrogen signals through its steroid nuclear receptor, in order to sustain the growth and proliferation of ER expressing cells. PR is a known downstream target of ER signalling, therefore epithelial cells with a functioning ER will generally express PR too. The ERα gene was cloned in 1986 [33] and sequence comparisons highlighted the similarity between ER and the glucocorticoid receptor (GR) [34]. This helped to define the family of transcription factors, known as the nuclear receptor (NR) superfamily. The NR superfamily has grown to become one of the largest families of transcription factors, with 48 members in humans, and includes receptors for androgen (AR), progesterone, vitamin D3 (VDR), thyroid hormone (TR), retinoids (RAR), fatty acids, bile acids, prostaglandins, xenobiotics, phospholipids, and heme, as well as "orphan" receptors that have no known ligand or that may not require ligand binding for their activities [35, 36]. The members of this family share a structural homology and key functional domains.

The ER has a well described structural domain termed A/B through to F (Figure 1.5). The A/B domain is located at the N-terminus and contains an activation region known as AF1 which is responsible for constitutive and ligand independent transcriptional activity of ER. The C domain encompasses the DNA-binding domain, which is responsible for specific DNA binding and receptor dimerisation. The D domain consists of a flexible hinge between the C and E domains and contains a nuclear localisation signal. The E domain is referred to as the ligand binding domain (LBD) and is involved in ligand binding and receptor dimerisation. It also harbours a second nuclear localisation signal and the activation function 2 (AF2) which is responsible for ligand dependent activation of ER. Finally the F domain lies at the extreme C-terminus and whilst it is unnecessary for transcriptional activation of the ER, it is proposed to exert a modulatory effect on the activity of AF1 and AF2 [37, 38].

Estrogen and estrogen-related molecules can bind to two distinct receptors; ERα and ERβ which are encoded by two separate genes, located on chromosome 6 and 14, respectively [39]. ERα and ERβ, each have a relatively high tissue-specific expression with overlapping distribution in some tissues. ERα is predominantly expressed in the pituitary gland, ovaries, uterus, liver, kidneys, adrenals and the mammary gland whilst ERβ is expressed in prostate, bone, the granulosa cells of the ovaries and the lungs [39-41]. Normally, when bound to their ligand, the ERs function as transcription factors of specific target genes. Although ERα and ERβ share a high structural homology, the two isoforms have less than 20% amino acid identity in the A/B and F domains which could account for the subtype specific actions on

target genes [41]. ERβ has a relatively weak AF1 activity and shows a differing response to agonist/antagonist ligands [42].

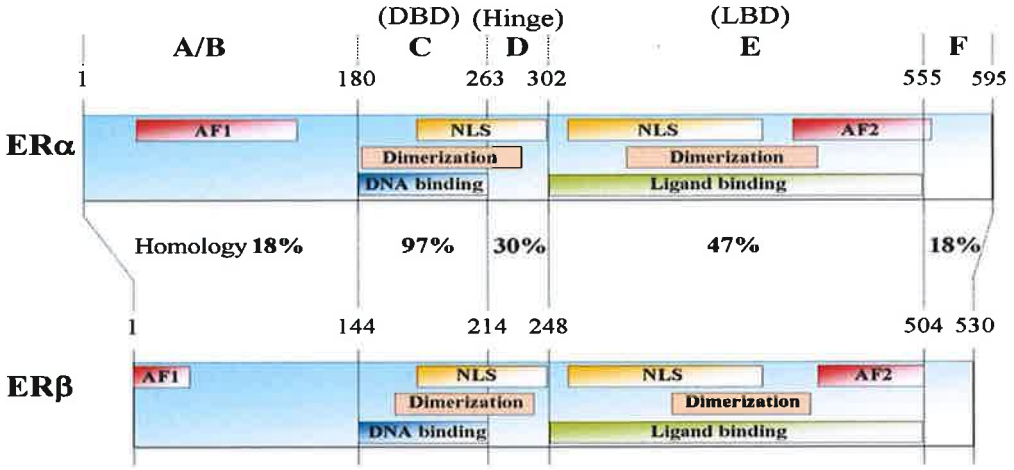


Figure 1.5. ERα and ERβ nuclear receptor structure. Both ERα and ERβ are characterized by a well-defined domain organization: the A/B domain at the N-terminus contains the ligand-independent transcriptional activation function 1 (AF1), the C domain represents the DNA-binding domain (DBD) and contains a nuclear localization signal (NLS), the D domain corresponds to the hinge region, the E domain harbours the ligand-binding domain and the ligand-dependent transcriptional activation function 2 (AF2) and the F domain at the C-terminus. Numbers outside each box refer to amino-acid number. Percentage of amino-acid homology for each domain is also shown. *Adapted from Zilli, 2009* [39] .

With regards to the mammary gland, a deletion of ERβ has no adverse effects on ductal and alveolar development, whereas both stromal and epithelial ERα seem to be required for normal ductal elongation and outgrowth during puberty [43]. In the mammary gland, ERβ is thought to have pro-differentiative and anti-proliferative effects therefore it is the ERα isoform that is commonly associated with carcinogenesis and tumour progression in the breast [44].

1.2.4. Steroid Receptor Coregulator (SRC) Proteins

Coregulator proteins serve as a fine-tuning mechanism by increasing or reducing the transcriptional activity of the relevant NR. The significance of NR coregulator proteins came to light when *in vitro* experiments using purified NRs and basal transcription factors proved relatively incapable of inducing transcriptional activation on their own [45, 46]. Furthermore, NRs were also shown to compete with each other for these essential coregulators as overexpression of one NR appeared to inhibit the transactivation function of another [47].

The p160 SRC family of coactivators have been well researched and validated with regards to their role in NR transcriptional regulation. The members include SRC-1 (NCOA1), SRC-2 (NCOA2, TIF2 or GRIP1) and SRC-3 (AIB1, p/CIP, ACTR, RAC3 or NCOA3) [18, 19]. Each member is approximately 160 kDa in size and their sequences are largely conserved across family members and also across species. The p160 SRC family members also have overlapping coactivator functions and transfection assays have shown that all three can coactivate ERα, PR and glucocorticoid receptor (GR) [48]. The potential for functional redundancy among the three members may serve to ensure a safety mechanism in the regulation of numerous important biological processes that are associated with the various NR signalling pathways. NR coactivators are unable to bind directly to the DNA. Instead they form multiple contacts with the NR and with each other in multi-protein cooperative coactivator complexes. Initial investigations into co-activator complexes reported that steady-state SRC complexes consist of six to ten stably associated proteins and many more loosely-bound proteins [49]. The versatile structural domains of SRC-1 and the other SRC family members grant them a central position in such complexes, from which they regulate multiple biochemical processes critical for the successful execution of transcription. Recently, the SRC proteins have been reclassified as "master genes" on the basis of their ability to bind across unrelated families of transcription factors and coordinate the regulation of multiple genes in multiple complex physiological states [50].

For genes which are downregulated in response to NR signalling, transcriptional corepressors are commonly recruited to the promoter site by the relevant transcription factors. Corepressors function in complete contrast to the coactivators and function to alter the chromatin towards an inactive state. The corepressors SMRT (silencing mediator of retinoid and thyroid receptors) and NCoR (nuclear receptor corepressor) recruit and activate histone deacetylases which retain the chromatin in a condensed configuration so that no transcriptional activity can occur [51]. There have been some instances of note where the roles of coactivator and corepressor appeared interchangeable. Specifically the coactivator SRC-2 was shown to function as a corepressor at the ERα-repressed TNFα promoter [52], and the corepressor SMRT coactivated TRα-driven transcription at a negative thyroid response element [53] It has been suggested that the promoter context of a transcriptional reaction complex can influence the functionality of a coregulator protein. For example, SRC-2 coactivates GR-directed transcription when GR is bound directly to glucocorticoid response elements, but it represses GR directed transcriptional when GR is binding indirectly to the DNA via activating protein 1 (AP-1) [51, 54].

1.3. Steroid Receptor Coactivator 1 (SRC-1)

SRC-1 is the founding member of the p160 family. SRC-1 was first discovered in 1995 in a yeast two-hybrid screen based on its interaction with the ligand binding domain (LBD) of progesterone receptor (PR) [48]. This work represented the first cloning of an authentic NR coactivator. SRC-1 demonstrated the ability to interact with and coactivate numerous NRs in the presence of hormones. These SRC-1 coregulated NRs include PR, GR, ER, TR, RXR, hepatocyte nuclear factor 4 (HNF4α) and peroxisome proliferator-activated receptor γ (PPARγ) [48, 55, 56]. The binding affinity of SRC-1 for these NRs has been shown to vary depending on where it specifically binds the NR. SRC-1 can bind NRs via its central region or less commonly via its C-terminal domain. The central domain of SRC-1 has been shown to be unable to bind to the AR and only exhibits a poor binding affinity for GR. In contrast, the C-terminus of SRC-1 exhibits a poor binding affinity for ER, VDR, RAR and TR, relative to its central domain [57]. Importantly, SRC-1 coactivator activity is not limited to the transcriptional coactivation of NRs, SRC-1 is also capable of coactivating other non steroidal transcription factors such as AP-1, serum response factor, NFκβ, Ets2, PEA3, HOXC11, ITGA5 and Twist [58-64].

1.3.1. Structural and functional domains of SRC-1

Just as with the ER, the SRC-1 protein structure is composed of several distinct functional domains. The N-terminus contains a basic helix-loop-helix-Per/Ah receptor nuclear translocation/Sim (bHLH/PAS) motif and is the most conserved region among the SRC family members with 75% similarity [65]. The bHLH/PAS domain is important for the protein-protein interactions that recruit secondary coactivators or co-coactivators to maximize the transcriptional activity of NRs (Figure 1.6). The domain is also important for the dimerisation of SRC proteins and for the differential regulation of target genes [66, 67]. The central region of the SRC-1 protein contains the nuclear receptor interaction domain (NRID). This domain contains three α-helical LXXLL (L, leucine; X, any amino acid) motifs which are essential for interaction of SRC-1 with its specific NRs (Figure 1.6) [68, 69]. These LXXLL motifs are the most common feature among the highly diverse group of coregulators [70]. Distinct LXXLL motifs along with their specific flanking sequences exhibit different binding affinities for different NRs inferring that NRs may prefer one LXXLL motif over another in the same coactivator or even prefer one coactivator over another [70, 71].

Secondary structural analysis of the central region shows that LXXLL motifs form an amphipathic α-helix which binds to a hydrophobic cleft that forms in the ligand-binding

domains of NRs once ligands have bound [72]. Interactions between the leucine residues of the LXXLL motif and the hydrophobic cleft stabilize the SRC-1/NR complex [73]. Mutation of these leucine residues have been shown to inhibit the binding of SRC-1 to the ERα LBD *in vitro* and SRC-1-mediated activation of ERα *in vivo* [69].

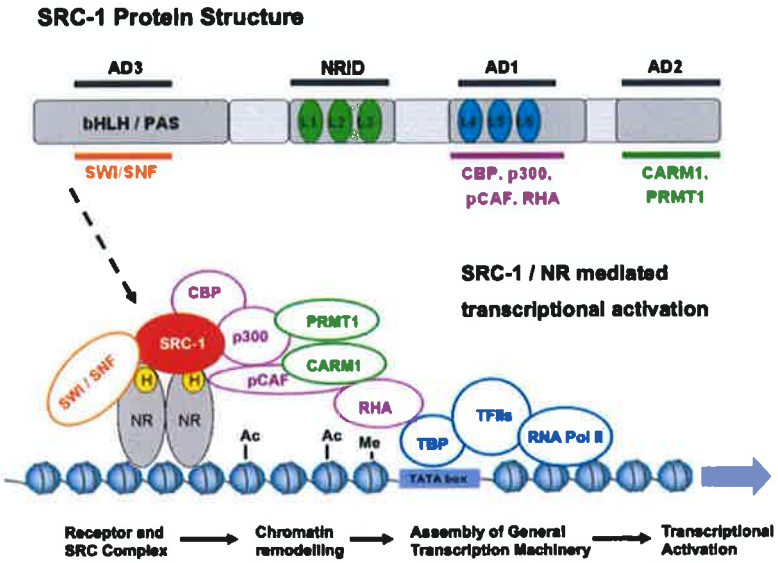


Figure 1.6. SRC-1-mediated coactivation of NRs. To exert its coactivation function in transcription, SRC-1 interacts with hormone (H)-bound nuclear receptors (NRs) to recruit other components of a large coactivator complex to the hormone response elements of a target gene. Specifically, SRC-1 binds NRs through one of its three LXXLL motifs (L1, L2 and L3) in the NR interaction domain (NRID) and interacts with CBP and p300 through its activation domain 1 (AD1), with CARM1 and PRMT1 through its AD2 and with SWI/SNF through its AD3. p/CAF is a p300/CBP-associated factor. CBP, p300 and p/CAF are histone acetyltransferases. CARM1 and PRMT1 are histone methyltransferases. RHA is a RNA helicase. SWI/SNF is an ATP-dependent chromatin remodelling complex. The formation of such a coactivator complex results in chromatin remodelling and bridges the hormone-activated NRs with the general transcription machinery for transcriptional activation of their specific target genes. bHLH/PAS: basic helix-loop-helix-Per/Ah receptor nuclear translocation/Sim motif; Ac: acetylation; Me: methylation; TBP, TATA-binding protein. *Adapted from Walsh et al, 2012 [74].*

1.3.2. Activation domains of SRC-1

The C terminal domain of SRC-1 contains two intrinsic transactivation domains: AD1 and AD2. The AD1 domain is required for the recruitment of secondary coactivators, such as cAMP response element binding protein (CREB)-binding protein (CBP) and the histone acetyltransferase p300 (Figure 1.6). Once SRC-1 has complexed with its ligand-bound receptor at the DNA, it must bring in additional proteins such as CBP and p300 to acetylate

histone residues within the enhancer and promoter regions of the target gene so that transcription can successfully occur [75]. Furthermore, observations from the p300 protein structure have confirmed that it is the SRC-1 interacting domain and not the ERα-interacting domain, which is essential for p300-mediated coactivation of ERα [75]. Additional experiments have demonstrated that the C terminus of SRC-1 and another p160 family member, AIB1, have intrinsic histone acetyltransferase activity (HAT); however, it remains unclear if such HAT activity is significant for target gene activation as the SRC-1 intrinsic HAT activity is much weaker than that in the CBP/p300 proteins [76, 77]. SRC-1 also uses its AD1 domain to interact directly or indirectly via CBP/p300 with another HAT protein known as p/CAF (Figure 1.6). p/CAF primarily acetylates histone H3 and H4 to further facilitate chromatin remodelling at the site of NR target genes [78]. The AD1 domain is also responsible for interactions with the general transcription machinery [79, 80] The AD1 has three LXXLL-like motifs and if any of these motifs are mutated there is a consequential disruption of SRC-1's ability to functionally interact with the necessary transcriptional components [81, 82]. Additionally, RNA helicase A (RHA) may be recruited to the SRC-1–CBP/p300 complex at the AD1 domain and interacts with RNA polymerase II, bringing NRs further into touch with the assembling transcriptional complex (Figure 1.6) [83].

The AD2 domain of SRC-1 recruits histone methyltransferases (HMTs) such as coactivator-associated arginine methyltransferase 1 (CARM1) and protein arginine N-methyltransferase 1 (PRMT1) (Figure 1.6). CARM1 is a histone H3 specific arginine methyltransferase, which specifically methylates histone H3 at arginines 2, 17, and 26. CARM1 can only enhance NR transcription in the presence of a SRC protein and mutation of its binding domain will reduce its HMT activity and ability to act as a co-coactivator [84].

Recently, a third activation domain (AD3) has been identified in the bHLH/PAS domain. This region also binds co-coactivators for NR mediated transcription. Many of these AD3 co-coactivators function synergistically with the other co-coactivators that bind to the AD1 and AD2 domains of SRCs [51].

1.3.3. Normal SRC-1 function in endocrine tissues

SRC-1 knockout mice were generated to investigate the normal biological functions of SRC-1 in various tissues. The targeted deletion of the SRC-1 gene in mice disrupted its nuclear receptor binding and transcriptional activation functions as well as its recruited HAT activity from the secondary coactivators such as CBP, p300 and p/CAF [85]. Despite these disruptions, SRC-1 null mice exhibited no obvious phenotype. Both male and female homozygotes were fertile and displayed a similar growth rate compared to the wild type

mice. However, further investigations revealed that steroid activity was partially impaired in SRC-1 null mice. Disruption of SRC-1 reduced estrogen-induced uterine growth; hence SRC-1 null mice achieved less than 60% uterine growth compared to that of the wild type mice when treated with estrogen [85]. Furthermore, SRC-1 is required for estrogen and progesterone-induced activities of ER and PR in the uterus but not in the breast [86].

SRC-1 is also expressed in the glandular and stromal cells of the normal endometrium. Immunoprecipitation assays revealed that estrogen-induced SRC-1 interaction with endometrial ERα only occurred during the proliferative phase of the menstrual cycle, indicating that SRC-1 has a regulatory role specific to this stage of the cycle [87, 88].

SRC-1 is also required for normal mammary duct elongation during puberty and alveolar development during pregnancy. Alveoli number and size were reduced in SRC-1 null mice compared with wild type mice of the same pregnant stage, although significantly the null mice did retain the capacity to produce milk [85].

Loss of SRC-1 activity also causes partial resistance to thyroid hormone signalling [80, 89] and has a role to play in the maintenance of bone mineral density by sex hormones, which is necessary for protection against osteopenia and osteoporosis [90].

1.4. Endocrine Signalling in Breast Cancer

As previously mentioned, the ER signalling network is highly complex and capable of modulating the expression of hundreds of genes, both by upregulation and by downregulation. In a normal healthy breast, only 15–25% of mammary epithelial cells express ERα however in a primary breast cancer tumour, this number increases to between 70% and 80% [91]. Whilst the alteration in ERα function is not yet fully understood, the importance of estrogen dependence and the ERα signalling pathway for growth and survival of tumour cells is very apparent. Decades of scientific research have been invested in elucidating these corrupted signalling mechanisms and the predominant consensus suggests that ERα signalling in breast cancer is mediated through two potential signalling pathways; the genomic nuclear signalling pathway and the non-genomic membrane signalling pathway. To date, the genomic signalling ERα has been the most cited of the two pathways however momentum is growing with regards to the tumourigenic potential of the non genomic ERα activity and its potential for receptor crosstalk with the various growth factor pathways.

1.4.1. Genomic ERα signalling pathways

In the genomic pathway, estrogen can easily diffuse through the phospholipid bilayer of the cell membrane and into the cytoplasm. Once in the cytoplasm, estrogen binds to the LBD of the intracellular ERα which induces a conformational change in the receptor structure. ERα becomes phosphorylated and dissociates from its chaperone proteins to homodimerise or form ERα/ERβ heterodimers [39].

Once the estrogen-ERα complex is activated, it can then translocate to the nucleus. In the nucleus, the estrogen-ERα complex is able to interact with the genome directly via specific estrogen response elements (EREs) or indirectly via non classical promoter elements such as AP-1, SP1 or NF-κβ sites or also via Jun/Fos proteins (Figure 1.7). Once bound in this manner, ERα with its associated coactivators or corepressors can form sophisticated transcriptional coregulatory complexes which will modify the histones and remodel the chromatin in order to then activate or silence the transcription of a specific target gene [38]. Interestingly, genome wide analysis of ERα transcription factor binding sites in breast cancer cell lines reported that functional ERα binding sites can be located at a considerable distance from the transcriptional start site [92].

1.4.2. Non genomic ERα signalling pathways

Low levels of ERα have been found outside the nucleus in the cytoplasm and at the membrane. ERα localises to the cytoplasmic membrane by receptor palmitoylation and by binding to the scaffolding protein, caveolin1 [93]. Non genomic signalling refers to gene activation via either cytoplasmic or membranous ERα (Figure 1.7). Generally, the non genomic action of estrogen is associated with rapid ERα function whereby the effects are seen within minutes and appear to occur too quickly for a transcriptional reaction [39]. Once estrogen binds to the membrane ERα, the ERα forms a homodimer which directly interacts with adaptor proteins such as Src, p85 subunit of PI3K, and G-proteins. These complexes subsequently associate with growth factor receptors e.g. EGFR, IGFR, HER2, cytoplasmic kinases e.g. MAPKs, PI3K, AKT and the mammalian target of rapamycin (mTOR) signalling enzymes eg adenyl cyclase [94]. Together, they then phosphorylate various transcription factors (TFs) and coregulators, including components of the ERα pathway that enhance gene expression on EREs and other response elements [94-96]. Non genomic ERα signalling also effectively activates growth factor receptor signalling, including the PI3K/AKT and the Ras/p42,44 MAPK pathways and thus can alter the expression of genes normally regulated by growth factors [94, 97].

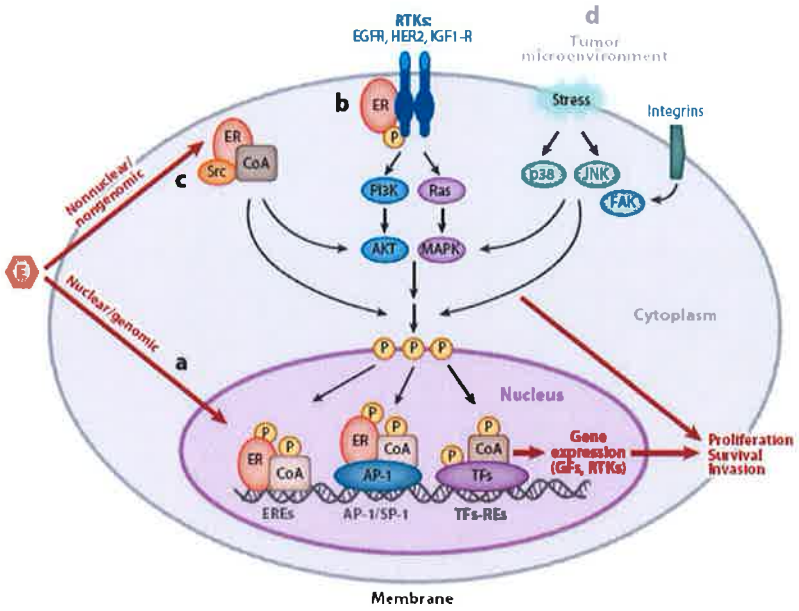


Figure 1.7. Mechanisms of estrogen receptor α (ERα) action in breast cancer. (a) Estrogen (E)-bound ERα, acting as a transcription factor in the nucleus (genomic activity), binds to DNA sequences in promoter regions of target genes either directly at EREs or indirectly via protein-protein interaction with other transcription factors at their specific DNA-responsive sites (e.g., AP-1 or SP-1). ERα can then recruit in coactivator complexes (CoA) to induce or modulate gene transcription, including genes encoding growth factors (GFs) and receptor tyrosine kinases (RTKs). (b) ERα localized outside the nucleus and/or at the cell membrane associates in response to estrogen with GF RTKs (e.g., EGFR, HER2, and IGF1-R) and (c) with additional signalling and coactivator molecules (e.g., the Src kinase). These complexes activate multiple downstream kinase pathways (e.g., SRC, PI3K/AKT, and Ras/p42/44 MAPK), which in turn phosphorylate various transcription factors (TFs) and coregulators to enhance ERα directed on EREs and other response elements (REs). (d) Signalling from the microenvironment activates stress-related pathways and members of the integrin family. These pathways then trigger downstream kinase pathways [e.g., FAK (focal adhesion kinase), JNK (c-Jun N-terminal kinase), and p38 MAPK] that can further modulate components of the transcriptional machinery, including ERα. *Adapted from Osborne and Schiff, 2011 [22]*

Another modulator of non-genomic ERα signalling is the tumour microenvironment which can activate stress-related pathways and members of the integrin family. These pathways then trigger downstream kinase pathways [e.g., FAK (focal adhesion kinase), JNK (c-Jun N-terminal kinase), and p38 MAPK] that can further modulate components of the transcriptional machinery, including ERα [94]. The non-genomic activity is highly regulated by coregulatory proteins, and it is influenced by the overall signal transduction pathways operating in a specific cell or tumour milieu [39]. As a tumour develops, the genomic and the non-genomic ERα signalling pathways can become complementary and even synergistic to provide breast tumour cells with proliferation, survival, and invasion stimuli [94].

1.4.3. ERα independent signalling pathways

As research into endocrine signalling mechanisms progresses, it has been noted that the ERα signalling pathway can also be regulated independently of estrogen activity by various forms of post-translational modifications. The most relevant of these modifications with regards to ERα activation is receptor phosphorylation which is generally elicited by kinase mediated cellular signalling events. Evidence has shown that membrane receptor tyrosine kinases (RTKs) can increase ERα phosphorylation and increase its transcriptional activity [98]. The RTKs are a group of high-affinity cell surface receptors for numerous growth factors, cytokines, and hormones. These proteins are key regulators of normal cellular processes but also have a well established role in the development and progression of many types of cancer. Approximately 20 classes of RTKs have been identified; these include the well known epidermal growth factor receptor (EGFR), the HER2 receptor as well as the insulin-like growth factor receptor (IGF1-R) [99]. Generally, when a growth factor binds to the extracellular domain of an RTK, it induces a dimerisation with other adjacent RTKs. Once dimerisation has occurred, it drives the rapid activation of the protein's cytoplasmic kinase domains. The receptor itself becomes autophosphorylated on multiple specific intracellular tyrosine residues. Phosphorylation of these membrane kinases activates downstream signalling pathways to phosphorylate ERα as well as its coactivators and corepressors at multiple sites to influence specific target gene expression [94, 99, 100].

Crosstalk between the growth factor receptor and ERα pathways has been established through several other mechanisms as well. Estrogen can increase the expression of RTK ligands such as transforming growth factor-α (TGFα), which can then activate the growth factor receptor pathway [94, 101]. Estrogen can also regulate expression of various RTKs although its regulation appears to be selective for different receptors; estrogen has been shown to decrease the expression of EGFR and HER2 whilst it increases the expression of IGF1-R [102, 103]. Furthermore, it has been demonstrated that activation of the PI3K/AKT and the p42/44 mitogen activated protein kinase (MAPK) pathways by these RTKs, can in turn, downregulates the expression of ERα and PR [104, 105].

The observed crosstalk and bidirectional signalling between the ERα and growth factor pathways is capable of activating the transcriptional function of ERα, as well as reducing estrogen dependence of tumour cells by downregulating the expression of ERα. These complex signalling mechanisms can be utilised and manipulated by breast cancer tumour cells in order to escape ERα targeted therapies [22].

1.4.4. The role of SRC-1 in breast cancer signalling pathways

SRC-1 protein expression is very low in normal human mammary gland ductal epithelial cells although it is steadily detectable in the nuclei of normal mouse mammary ductal epithelial cells. However in human breast tumours, SRC-1 protein expression was reported to be between 19% and 34%. This aberrant SRC-1 expression is significantly associated with large, high grade tumours, HER2 positivity, disease recurrence and resistance to endocrine therapy. SRC-1 expression also serves as an independent predictor of disease-free survival in clinical patient populations [61, 106, 107].

The role of SRC-1 in mammary tumour initiation, progression and metastasis has been investigated by knocking out SRC-1 in the *MMTV-PyMT* (mouse mammary tumour virus-polyoma middle T) transgenic mice [108]. Analysis of mammary tumourigenesis in *SRC-1*$^{-/-}$*;MMTV-PyMT* and in *wildtype MMTV-PyMT* mice revealed that SRC-1 is not required for mammary tumour initiation and growth but rather its activity is specific with mammary tumour cell intravasation and metastasis to the lungs. Blood samples taken from the *wildtype MMTV-PyMT* mice showed a significantly higher number of mammary tumour cells in circulation than the samples taken from the *SRC-1*$^{-/-}$*;MMTV-PyMT* mice, indicating that SRC-1 promotes migration of breast cancer cells and invasion of these tumour cells into the vascular system.

SRC-1 expression in the mammary tumour cells also upregulates the expression and secretion of colony stimulating factor-1 (CSF-1) to facilitate the recruitment of macrophages to the mammary tumour sites [108]. Macrophage recruitment also has well established associations with invasiveness and metastasis of breast tumour cells to local and distant sites [109]. At the tumour site, macrophages secrete factors such as EGF and TGF-β1 that stimulate proliferation and invasiveness of tumour cells [110]. Macrophages also induce angiogenesis via secretion of proangiogenic cytokines such as TNF-α and IL-8. Importantly, the metastatic nature of SRC-1 is intrinsic to the tumour cells as demonstrated by reciprocal transplantation of the *wildtype MMTV-PyMT* and *SRC-1*$^{-/-}$*;MMTV-PyMT* tumours into *SRC-1*$^{-/-}$ and *wild type* recipient mice [108].

In the *MMTV-PyMT* model, loss of SRC-1 also correlated closely with reduced expression of the HER2 oncogene as well as with reduced levels of the steroid independent transcription factor Ets2, thus further confirming the associations between SRC-1 and growth factor signalling effectors in breast cancer [108]. SRC-1 also interacts with Ets2 in HER2 positive breast cancers [106, 107, 111]. Under the influence of HER2, Ets2 and SRC-1 are recruited to

the promoter of the *c-myc* oncogene to facilitate its transcription and to promote breast tumour cell survival, metastasis and therapeutic resistance [112, 113].

A second SRC-1 knockout transgenic model; the *SRC-1*$^{-/-}$;*MMTV-neu* mouse breast cancer model was generated by crossing *MMTV-neu* transgenic mice with SRC-1 knockout mice. In this model, SRC-1 deficiency increased tumour latency and reduced tumour cell proliferation index and lung metastasis. As observed in the *SRC-1*$^{-/-}$;*MMTV-PyMT* model, the *wildtype SRC-1*$^{-/-}$;*MMTV-neu* tumours were also more differentiated than the *MMTV-neu* tumours [114]. The findings from both *SRC-1*$^{-/-}$;*MMTV-PyMT* and *SRC-1*$^{-/-}$;*MMTV-neu* breast cancer models indicate that SRC-1 plays a crucial role in promotion of breast cancer metastasis. Observations from the *MMTV-neu* model did suggest a role for SRC-1 in the enhancement of tumour initiation and growth, hence specific SRC-1 activity in breast cancer tumour cells could be relative to the individual oncogene-induced tumourigenic pathways. In addition, since the tumours from both models are ER negative, SRC-1 likely promotes tumourigenesis and metastasis in these models through estrogen-independent pathways.

1.4.5. A specific role for SRC-1 in metastatic breast cancer

Two distinct mechanisms have been elucidated for SRC-1 in metastatic breast cancer. Firstly, SRC-1 impacts on the epithelial to mesenchymal transition (EMT) and epithelial polarity of the tumour cells via regulation of the EMT transcription factor Twist. EMT is characterized by the loss of epithelial differentiation and the gain of mesenchymal properties within the cellular population. This transition enables tumour cells to invade and survive in the stromal tissues. In essence, EMT is considered to be an early step towards metastasis and is characterized by certain gene expression changes [64, 115]. *In vitro* experiments from the wild type and SRC-1 knockout cell lines show that SRC-1 inversely correlates with the major EMT hallmark protein, E-cadherin, by coactivating the EMT transcription factor, Twist, which consequently is a negative regulator of E-cadherin [116]. SRC-1's regulation of Twist expression appears to involve estrogen independent signalling mechanisms. It has been shown that Twist protein expression is under the control of the MAPK pathway in HER2 overexpressing cancer cells. SRC-1 has been reported to bind with the downstream MAPK effector; polyoma enhancer activator 3 (PEA3), to coactivate Twist transcription [64]. The SRC-1 mediated upregulation of Twist in HER2 positive tumours correlates with negative PR status in the tumour cells. Since PR is a transcriptional target of ERα and since PR negative cancers are associated with a more aggressive basal phenotype, this infers that SRC-1 can work independently of estrogen to progress a tumour towards an increasingly aggressive and metastatic state [117].

A second mechanism of SRC-1 action in metastatic breast cancer involves the SRC-1 mediated upregulation of integrin α5 to promote cell migration and invasion. In this instance, SRC-1 can assist metastatic activity via the non-steroidal transcription factor AP-1 and the tumour microenvironment. Resident fibroblasts produce abundant extracellular matrix (ECM) proteins to provide anchorage for tumour cell adhesion and migration [118]. Heterodimeric integrin transmembrane receptors bind to ECM proteins to transport signals bidirectionally across the cell membrane, allowing cells to respond to environmental changes [118]. In breast cancer, these integrin receptors have been associated with tumour cell survival, growth and metastasis in an anchorage independent manner. In particular, the heterodimeric receptor, α5β1, is increased in malignant breast cancer and is associated with poor prognosis [119]. SRC-1 and AP-1 complex together at the promoter of the integrin α5 gene to enhances its transcription. In the absence of SRC-1, cell adhesion and migration via the ECM-integrin-focal adhesion kinase (FAK) pathway is markedly reduced. The downstream signalling components of this pathway are also significantly reduced in the absence of SRC-1 and the integrin α5 [63], suggesting that integrin α5 is a key target gene of SRC-1 in its efforts to orchestrate the local invasion and metastatic survival of breast cancer tumour cells.

1.5. Endocrine Treatment of Breast Cancer

1.5.1. Selective ER downregulators (SERDs)

The selective ER downregulators (SERDs) are a class of anti-estrogenic molecules which are used for the treatment of breast cancer. ICI182,780 (fulvestrant), was the first of this class of molecules to be clinically approved; it is a 7-alkylsulphinyl analogue of 17β-estradiol which is anti-proliferative and inhibits ERα-mediated gene transcription via degradation of the ERα in breast cancer cells [120]. Fulvestrant has a unique C7 side chain in its chemical structure therefore when fulvestrant binds the ERα LBD, it induces a distinct conformational change in the structure of ERα, which then disrupts the ability of receptor to dimerise and bind to the DNA. Fulvestrant also impacts on the ability of ERα to translocate to the nucleus which then leads to increased degradation of ERα in the cytoplasm via the ubiquitin–proteasome pathway [121].

Studies on fulvestrant have shown its ability to effectively reduce ERα expression in breast cancer cells and also its ability to inhibit the growth of Tamoxifen-resistant breast tumour xenografts [122, 123]. However, these promising observations failed to translate to the

clinic where the therapeutic results to date have been quite disappointing [124]. Initially, the lack of efficacy was attributed to differing mechanisms of tamoxifen resistance in animals versus human tumours. However, it now appears from the results of additional studies that fulvestrant has extremely poor bioavailability therefore it is difficult to get high enough levels of the drug at the tumour site to significantly affect the quantitative turnover of ERα. Indeed, a sequential tumour biopsy study has indicated that even after long-term fulvestrant treatment at the approved dose, ERα is still present at approximately 50% of the original baseline[125]. Results of recent trials comparing higher doses of fulvestrant to that currently approved demonstrate increased serum steady state levels and correspondingly improved response rate, as well as increased ERα turnover [124, 126, 127]. Although the pharmaceutical properties of fulvestrant continue to limit its use, there is a tremendous amount of interest in developing SERDs with improved pharmacodynamics [128].

1.5.2. Aromatase Inhibitors (AIs)

As previously mentioned in section 1.2.2., estrogen can be produced from sites other then the ovaries via peripheral aromatization of androgens. The aromatase enzyme is a member of the cytochrome P-450 superfamily and the product of the CYP19 gene [129]. Aromatase is highly expressed in the granulosa cells of ovaries, and its expression fluctuates in response to cyclical gonadotropin levels. Aromatase, is also present, at lower levels, in subcutaneous adipose tissue, liver, muscle, brain, normal breast, and breast-cancer tissue [130, 131]. Aromatase converts androstenedione and testosterone produced by the adrenal gland to estrone and estradiol, respectively. AIs work to block the aromatase enzyme which results in a severe decrease in plasma estrogen levels.

This effect is more relevant in post-menopausal women in whom estrogen production occurs essentially by peripheral aromatization of androgens. During the menopause, estradiol levels in the plasma fall from approximately 110 pg/mL to low but stable levels of 7 pg/mL [131]. Despite this dramatic reduction in circulating estradiol, post menopausal women can have quite high levels of estradiol expressed within a breast cancer tumour and this observation has been somewhat attributed to the presence of aromatase within the tumour itself [132]. The use of aromatase inhibitors has largely been restricted to the treatment of postmenopausal women as its administration in premenopausal women, causes an increase in gonadotropin secretion and so must be given in conjunction with a gonadotropin inhibitor to suppress ovarian function [133].

Aromatase inhibitors were first described for clinical use in the late 1970s. Today, a third generation of aromatase inhibitors are in production with two classes in clinical use: the

steroidal AIs that binds the enzyme irreversibly, and the non-steroidal AIs which bind the enzyme reversibly [131]. These third-generation AIs have greater potency, specificity and reduced side effects compared with first and second-generation compounds, and therefore are increasingly used in the hormonal treatment of post-menopausal women with ERα positive breast cancer.

1.5.3. Selective ER Modulators (SERMs) – Tamoxifen

SERMs such as tamoxifen, raloxifene, and faslodex are a class of compounds endowed with estrogen agonistic and antagonistic activity in different tissues (Figure 1.8). SERMs competitively inhibit binding of estrogen to ERα and affect both the genomic and non genomic activity of ERα [39]. Tamoxifen is the most prominent SERM in clinical use today and its success is largely attributed to its ability to act as an antagonist in the breast to curb ERα-mediated proliferation whilst retaining agonistic activity in the bone and the cardiovascular system where the loss of estrogen can otherwise cause harmful side effects such as osteoporosis and atherosclerosis. Tamoxifen was first described in 1966 [134]and was initially licensed for use in women with advanced breast cancer in 1974. Since then, Tamoxifen has progressed to become the first-line adjuvant therapy for the treatment of women with ERα positive breast cancer. Furthermore, the World Health Organisation (WHO) lists tamoxifen as an essential drug for the treatment of breast cancer and it is estimated that more than 400,000 women are alive today as a result of Tamoxifen therapy, and that millions more have benefited from palliation and extended disease-free survival [135].

Tamoxifen is a nonsteroidal triphenylethylene agent, which was originally developed as a postcoital contraceptive [136]. However it was subsequently found to induce ovulation in women and was initially developed for use in fertility treatment. Important work from VC Jordan in the 1970s presented evidence to support a targeted role for Tamoxifen at ERα through the use of the 7,12-dimethylbenz[α]anthracene (DMBA)-induced rat mammary carcinogen model. Observations of Tamoxifen action in this model showed that breast cancer occurrence could be delayed by giving Tamoxifen after a carcinogenic insult with DMBA. The model also showed that breast cancer could be prevented if Tamoxifen is given at the same time as the carcinogenic insult [135]. These findings were the basis for future clinical trials looking at Tamoxifen as an adjuvant therapy and as a means of breast cancer prevention.

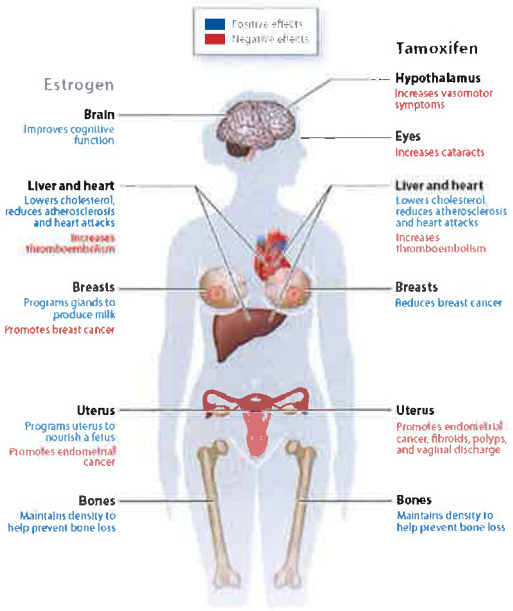


Figure 1.8. Positive and negative effects of estrogen and Tamoxifen action in the body. In adult women, estrogen and Tamoxifen show positive (protective) and negative tissue-selective effects. The development of new selective estrogen receptor modulators (SERMs) is aimed at maximizing the positive effects of estrogen but minimizing the negative effects. *Adapted from Coombes et al, 2011 [137].*

1.5.4. Mechanism of action for Tamoxifen

Tamoxifen is administered as tamoxifen citrate and undergoes extensive metabolism in the gastrointestinal tract and the liver, particularly by hydroxylation to generate 4-hydroxytamoxifen (4-OHT) and its secondary metabolite; endoxifen. Member of the cytochrome P-450 superfamily, specifically the enzyme, CYP2D6, catalyse the biotransformation of Tamoxifen into active and inactive metabolites [138].

As outlined in section 1.2.3, there are two 'activation' domains within the ERα protein structure, which facilitate the interaction of ERα with the transcription apparatus; AF1 and AF2. Any agent which induces a conformational change at these two activation domains changes the presentation of these domains to the transcription apparatus and ultimately impacts on the expression of their target genes [37]. It has been found that, in all contexts where AF-2 is required for ERα transcriptional activity, tamoxifen functions as a pure antagonist [139]. Tamoxifen binds to ERα which leads to a repositioning of helix 12 which causes a partial occlusion of the coactivator binding groove. This conformational change specifically disrupts the ability of the AF2 to interact with associated coactivators and instead promotes the recruitment of corepressors to the DNA site [39, 140]. Conversely, in tissues where AF1 activity is strong, Tamoxifen is predominantly agonistic in its effects. It

remains to be fully understood how Tamoxifen activates AF1 but there is evidence to indicate that phosphorylation of ERα and recruitment of coactivators similar to estrogen could be involved [141]. Specifically, it was proposed that, in tissues such as breast, where tamoxifen exhibited pure antagonist activity, AF-2 was required for transcriptional activity. On the other hand, in the uterus, bone and the cardiovascular system, where tamoxifen exhibits estrogenic activity, it was concluded that AF-1 activity was sufficient (Figure 1.8) [142].

Initially the mechanism of Tamoxifen anti-estrogenic activity was oversimplified. *In vitro* crystallisation studies of the ERα structure have shown that ERα is not simply a rigid molecule that exists in either an active or inactive state, but instead its conformation is determined by the nature of the bound ligand. Furthermore, Tamoxifen does not simply function passively as a competitive antagonist of ERα which freezes it in an inactive conformation. Instead Tamoxifen is now recognised as an active antagonist which has the ability to structurally alter the ERα, in order to interfere with the ERα-estradiol complex at multiple steps in the signal transduction pathway [142].

The AF domains are not the only molecular determinants which influence Tamoxifen SERM activity in various tissues. Studies have shown that the cellular availability of steroid receptor coactivators and corepressors can also impact on whether Tamoxifen behaves as an agonist or as an antagonist [143]. High levels of SRC-1 expression have been associated with agonistic effects of Tamoxifen in several endometrial cell lines. Overexpression of SRC-1 in breast cell lines induces an agonistic response to Tamoxifen as opposed to the expected antagonistic response. Furthermore, if SRC-1 expression is ablated in endometrial cells, Tamoxifen-agonistic activity is lost [144]. It has been observed that SRC-1-mediated agonistic effect of tamoxifen involves genes whose promoters do not contain a classical ERE [144] and depends largely on a direct interaction of SRC-1 with AF1 [145].

A recent study by Kressler et al, 2007, demonstrated a cooperative mechanism between the steroid receptor coactivators and the AF domains to mediate Tamoxifen-agonistic activity. In a defined Gal4-based system, the steroid receptor coactivator, PGC-1β acts synergistically with SRC-1 in promoting the agonist activity of Tamoxifen [146]. It showed that Tamoxifen binding to the ERα induced a physical and functional interaction of the coactivator PGC-1β with AF2. Since PGC-1β can also bind SRC-1, the simultaneous interaction of SRC-1 with AF1 creates a molecular bridge between AF1 and AF2 [146]. In this way specific AF2 coactivators, such as PGC-1β, may contribute to the agonistic activity of tamoxifen.

Differential observations of Tamoxifen activity have been noted in ERα-pathways that are regulated in a classical manner with a distinct ERE or in non classical pathways whereby ERα signals through a non ERE site such as AP1 or SP1. In endometrial cells, which exhibit very active non classical signalling pathways, Tamoxifen is agonistic, whereas in ERα-positive breast cancer cells which show minimal non-classical pathway activity, Tamoxifen functions as a pure antagonist [39, 147].

1.5.5. Clinical trials involving Tamoxifen

In 2011, the Early Breast Cancer Trialists' Collaborative Group (EBCTCG) updated their meta-analyses of long term outcome in 21,457 women with early stage breast cancer. Data was collected from 20 randomised trials with 5 years of adjuvant Tamoxifen versus observation or placebo [148]. The median follow up of these trials at the time of the update was 13 years and the results showed the administration of Tamoxifen for up to 5 years reduced the recurrence rate substantially for the first 10 years, with no subsequent loss of the gains made during that first decade. Breast cancer mortality was reduced by about one third throughout the first 15 years and importantly the relapse curves did not converge after year 10 which means that 5 years of Tamoxifen can prevent a high proportion of recurrence and potentially cure many patients rather than simply delay an inevitable event. Furthermore, Tamoxifen remains an attractive option for premenopausal or perimenopausal women as at these ages, treatment carries little risk of endometrial cancer. The other important finding from the EBCTCG was that the only factor which predicted Tamoxifen efficacy was the presence of ERα. No benefit was detected if ERα was zero but Tamoxifen was beneficial at very low levels of cytosolic ERα and the benefits did increase with ERα measurements [148].

1.6. Mechanisms of Endocrine Resistance in Breast Cancer

Undoubtedly the development of Tamoxifen has been a pioneering event in the treatment of breast cancer and there has been tremendous success as a result of its clinical application. Nevertheless, despite its successes, resistance to Tamoxifen and to endocrine therapies in general is becoming an increasingly prevalent occurrence and presents a serious clinical dilemma for both clinicians and patients. Data from the EBCTCG meta-analysis recorded a recurrence rate of 15.1% at 5 years rising to 24.7% at 10 years post diagnosis in women treated with Tamoxifen for 5 years [149].

1.6.1. Loss of steroid receptor expression

The ERα signalling pathway is a complex network with many levels of control including extensive crosstalk with growth factor signalling pathways. The increasing complexity of tumour associated signalling mechanisms serves as an advantage for tumour cells which are seeking to escape the effects of targeted therapies and ultimately pave the way for the development of a resistant tumour phenotype. Resistance to therapy can be described as intrinsic or acquired. Generally speaking, acquired resistance is thought to occur more commonly especially as a tumour develops. However intrinsic resistance has been strongly associated with the lack of ERα expression. As reported in the EBCTCG update, the presence of ERα was predictive of Tamoxifen response therefore it follows that lack of ERα would cause or contribute to resistance [150]. Furthermore in a certain number of breast cancers, an absence of ERα gene expression has been in part due to the aberrant methylation of CpG islands in the 5' regulatory region of the ERα gene [151]. Additional evidence to support the epigenetic silencing of ERα came from studies whereby the loss of ERα expression associated with an increased activity of histone deacetylases (HDACs) which kept the chromatin in a transcriptionally quietened state [152]. Current evidence suggests that clinical treatment with a HDAC inhibitor and DNA methyltransferase-1 (DNMT1) inhibitor could restore Tamoxifen sensitivity in ERα negative breast cancers [153].

Loss of PR is also a common feature in the development of Tamoxifen resistance and is generally indicative of an alteration in tumour steroid receptor profile. PR is lost more frequently than ERα with intervening endocrine therapy, and with this loss the tumour becomes more aggressive and patients have a worsened disease course than patients who maintain PR expression after resistance to one endocrine therapy [154, 155]. Just as with ERα, loss of PR has been associated with increased growth factor signalling and upregulation of the PI3K pathway [156].

1.6.2. Upregulation of growth factor signalling and crosstalk

As previously outlined in section 1.4.3, receptors such as EGFR and HER2, whose expression is often inversely related to ER expression have also been associated with driving down ERα expression in breast cancer [157]. Modification of ERα status has also been proposed as one of the main mechanisms of acquired therapeutic resistance in breast cancer. Loss of ERα over time in a tumour occurs in approximately 20% of patients treated with endocrine therapy [158]. It remains to be seen whether loss of ERα is the result of a clonal selection of ERα negative tumour cells from a heterogeneous population of cells or if post translational

modifications such as proteasomal degradation could be at play in what was an originally ERα positive cell population [22].

The growth factor pathways have been well researched with regards to their involvement in endocrine resistance. These pathways are important escape routes for tumour cells that are trying to escape ER targeted therapy and require alternative signalling mechanisms for proliferation and survival. Pathways such as the HER RTK family and receptors for insulin/IGF1, fibroblast growth factor (FGF), and vascular endothelial growth factor (VEGF), as well as cellular Src, AKT, and stress-related kinases, have all been implicated as alternate signalling mechanisms in endocrine resistance [22].

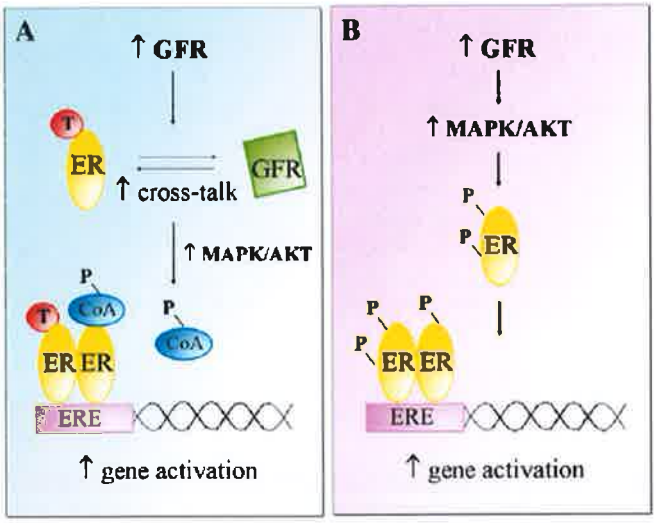


Figure 1.9. Overexpression of growth factor receptors (GFRs) causes endocrine resistance through different mechanisms. As resistance to therapy develops in a tumour cell, gene transcription occurs in conjunction with growth factor activity, initially, (A), in a ligand-dependent manner, creating a molecular milieu, e.g. phosphorylations (P) and activation of coactivators (CoA), in which tamoxifen (T) acts as agonist and, later, (B), in a ligand-independent manner, making critical phosphorylations (P) that enable ER to work in the absence of ligand (estrogen or tamoxifen). *Adapted from Zilli et al, 2009 [39].*

It has been suggested that treatment with Tamoxifen upregulates the expression of EGFR and HER2, which in turn signal through the MAPK and AKT pathways which then phosphorylate ERα at Ser118 and Ser167 and enable the continued expression of ERα target genes [159]. One such ER target gene, Amphiregulin, is also an EGFR ligand thus its resurgence under growth factor signalling further drives cell growth and proliferation and acts as a positive feedback loop for continuous proliferation within the cell [160]. Furthermore, activated membrane bound ERα can also activate HER2 signalling which serves to amplify the proliferative signal [94]. Tamoxifen can also activate membrane bound ERα and therefore also contributes to the upregulation of HER2/EGFR activity [161]. In

some instances of resistance, HER2 may become the dominant signalling pathway in a tumour and drive its progression in an entirely estrogen independent manner. Alternatively tumours progress and continue to alternate between ERα and HER2 pathways, with increasing cross talk and synergism and leaving clinicians and patients with very few options for successful targeted therapy.

1.6.3. Altered transcriptional coregulator expression

Finally, it has been shown that coregulators exert a huge amount of influence over ERα transcriptional targets. In certain circumstances, this influence extends further to determine therapeutic response. Overexpression of coactivators and/or down-regulation of corepressors have been shown to curtail the effects of endocrine therapy, especially in the case of SERMs [143]. Overexpression of the ERα coactivator AIB1 is associated with clinical and experimental tamoxifen resistance [162], and downregulation of the corepressor NCoR was documented in tamoxifen resistant experimental tumours [163].

1.6.4. A specific role for SRC-1 in endocrine resistant breast cancer

As well as a distinct role in the advancement of metastasis, SRC-1 has been specifically cited with regards to the development of resistance to endocrine therapy in breast cancer. Overexpression of SRC-1 has been reported to convert Tamoxifen from a transcriptional repressor to a transcriptional activator in breast cancer [164]. Similar effects have also been seen with NRs other than ERα, as overexpression of SRC-1 is also capable of inducing PR activity in the presence of the PR antagonist RU486 [165].

Conversely, treatment with Tamoxifen itself has been shown to increase the expression levels of SRC-1 in multiple studies [166, 167]. Furthermore, observations during the development of a Tamoxifen-resistant cell line reported that the expression of SRC-1, AIB1 and multiple cell cycle and cell adhesion genes was increased, while the expression of apoptotic genes and PR was concomitantly downregulated. Interestingly, the SRC-1 protein level was decreased in the resistant cells once Tamoxifen treatment was stopped for a 12 month period [168]. It was also reported that the co-association of SRC-1 and AIB1 with ERα was increased in LY2 endocrine resistant breast cancer cells following Tamoxifen treatment in comparison with endocrine sensitive MCF-7 cells. Moreover, the colocalization of SRC-1 and AIB1 with ERα was significantly enhanced in patients who have relapsed on endocrine treatment in comparison with those patients who did not undergo recurrence. Overall, analysis of the clinical data suggest that SRC-1 is a strong independent predictor of reduced

disease free survival and that the interactions of the p160 SRC proteins with ERα can predict the response of patients to endocrine therapy [169].

The relationship between SRC-1 and the growth factor pathways remains intertwined during endocrine resistance. In addition to serine 118 and 167, and in the presence of Tamoxifen, the HER2 pathway also induces protein kinase A (PKA) to phosphorylate ERα at serine 305. This phosphorylation alters the orientation of ERα and SRC-1 and switches Tamoxifen to an ERα agonist, resulting in RNA Polymerase II recruitment and transcriptional initiation [170]. These findings were supported by a separate study which confirmed the association between Tamoxifen resistance and this specific ERα phosphorylation [171].

Additional clinical data showed an increased co-association of SRC-1 with MAPK-activated transcription factors PEA3 and Ets2 in a Tamoxifen treated patient population [59, 110]. As expected, both of these co-associations correlated positively with HER2 positivity in the same population. In agreement with the upregulation of c-myc by Ets2 and SRC-1 in cultured cells, this upregulation was also observed in the Tamoxifen-treated patient population.

Recent data has shown that SRC-1 specifically interacts with the transcription factor MYB to directly regulate ADAM22, a non-protease member of the ADAM family of disintegrins [172]. Molecular analysis carried out in Tamoxifen-resistant cell lines discovered a role for ADAM22 in cellular migration and differentiation. In addition, expression of ADAM22 mRNA was increased in the Tamoxifen resistant tumours in xenograft mouse models. Furthermore, ADAM22 expression was identified by immunohistochemical analysis in a clinical patient population as an independent predictor of poor disease free survival [172].

1.7. SRC-1 and HOXC11 in endocrine resistance

HOXC11, a member of the homeobox superfamily of developmental proteins was identified as a non steroidal functional binding partner for SRC-1 in Tamoxifen-resistant LY2 cells [62]. This novel interaction between two proteins from two separate superfamilies presents many questions as to why these two proteins combine in endocrine resistance, what is their tumourigenic agenda and what is the potential impact on cell fate decisions, differentiation programs, and receptor switching?

1.7.1. The homeobox family

The HOX genes encode a family of transcription factors that play key roles in the determination and maintenance of cell fate and cell identity; they are 'master regulators' of normal development and control processes such as proliferation, apoptosis, migration and

invasion [173, 174]. In humans, 39 Hox genes have been identified and are found organised contiguously among four clusters. Hox genes within different clusters are arranged along the chromosomes 7, 12 and 2 as 13 paralogous groups [175]. The hallmark of the homeobox family of genes is that they share a common nucleotide sequence motif which encodes a 61 amino acid homeodomain. The homeodomain is a helix-turn-helix DNA-binding domain which mediates binding to a TAAT consensus site thus allowing these proteins to act as either activators or repressors of transcription [176]. Hox proteins can function as monomers or homodimers to directly drive the transcription of downstream targets It has also been suggested that Hox proteins can sequester other proteins to enhance or repress gene expression [177].

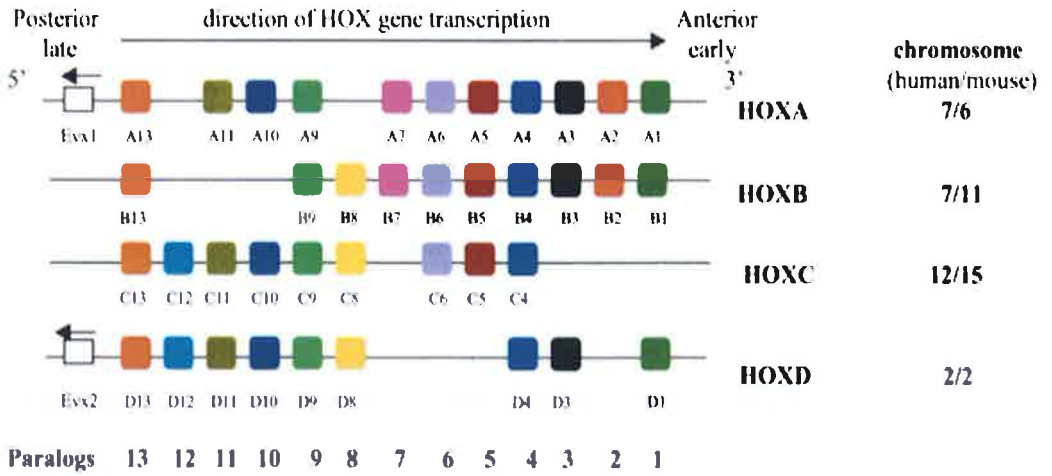


Figure 1.10. Organisation of the mammalian HOX clusters. *Adapted from Chen and Sukumar, 2003 [175]*

1.7.2. The role of Hox genes in the normal mammary gland

The first demonstrated role for homeobox genes in control of mammary function came from the targeted deletion of three paralogous *Hox* genes; *Hoxa9, Hoxb9* and *Hoxd9*. In triple mutant adult female mice, the mammary glands appeared normal however their development was severely impaired during pregnancy, leading to a failure in lactation [178]. Single mutant lines showed only a small decrease in newborn survival suggesting that functional cooperativity is at play between the three gene products during lobuloalveolar expansion. In contrast, *Hoxc6*-deficient mammary glands appeared underdeveloped at birth and revealed reduced ductal branching in the adult gland [179].

Five Hox genes were also identified in CID-9 mouse mammary epithelial cells. When these cells were cultured on reconstituted basement membrane in the presence of lactogenic hormones, they formed alveolar-like hollow spheres and differentiated to express milk proteins. *Hoxa1* and *Hoxb7* gene expression were downregulated by exogenous basement membrane suggesting that homeobox gene expression may be modulated *in vivo* by interactions between epithelial cells and components of the basement membrane. Such interactions have been shown to be critical for mammary differentiation and function and tend to be altered on neoplastic progression [173].

1.7.3. The role of HOX proteins in breast cancer

The accurate spatial and temporal regulation of Hox gene expression is imperative during embryogenesis to acquire the proper differentiated state and cell identity. There is an accumulating body of evidence that developmental genes are often misexpressed in human cancers and that this misexpression can impact neoplastic disease through the re-initiation of developmental programmes [180].

Misexpression of homeobox genes can lead to failure of differentiation or loss of the differentiated state, activation of anti-apoptotic pathways or adoption of an alternative cellular identity. This disrupted Hox gene expression can impact on many cell types within the body; causing affected cells to incur inappropriate characteristics such as loss of cell cycle, decreased apoptosis, altered cell-cell adhesions, altered hormone and growth factor responses or increased protease expression, all of which could easily contribute to cancer initiation or progression [176]. The expression of the Hox genes and their targets could be useful as a crude measure of the identity or functional state of a given cell thus Hox gene expression has become an area of interest in various cancer states.

HOXB7 has been identified as an ERα-responsive gene and overexpression of HOXB7 has been shown to render endocrine sensitive MCF-7 cells resistant to Tamoxifen via cross-talk between RTKs and ERα signalling. Extended treatment of MCF-7 cells with Tamoxifen resulted in progressively increasing levels of HOXB7 expression, along with EGFR and EGFR ligands. The study also showed up-regulation of EGFR occurs through direct binding of HOXB7 to the EGFR promoter, thus enhancing transcriptional activity. Finally, higher expression levels of HOXB7 in the tumour significantly correlated with poor disease-free survival in ERα-positive breast cancer patients on adjuvant Tamoxifen monotherapy [181].

A large study by Ma et al, used microarray analyses of ERα positive tumours from patients with disease recurrences despite receiving ongoing Tamoxifen treatment and found that HOXB13 expression was markedly upregulated compared to tumour samples from patients

without recurrences [182]. Later work showed that these patients who overexpressed HOXB13 were also at risk of distant metastasis at recurrence [183]. Another overexpression study using HOXB13 demonstrated that HOXB13 could suppress the ERα-dependent transcription and expression of components of the ERα pathway such as PR and cyclin D1. In such cases, there was no mutation of HOXB13 or its promoter or amplification of the gene itself, which suggests that the mutation may lie with the gene regulating HOXB13 or in a gene that leads to the activation of another growth factor pathway [177].

It is also hypothesised that aberrant Hox expression could impact on apoptosis to promote tumourigenesis. Loss of HOXA5 expression and methylation of the HOXA5 promoter were shown to correlate with reduced expression of the tumour suppressor p53, in both primary breast tumours and breast cancer cell lines [184]. Further *in vitro* studies identified a HOXA5 binding site on the p53 promoter whereby upregulation of HOXA5 induced p53 expression and increased the rates of apoptosis in the MCF-7 cancer cell line. HOXA10 was shown to have a similar effect on p53 in ERα positive breast cancer cell lines [185].

Whilst there is a general consensus as to the overall function of the Hox gene family, there is still limited information on the specific normal and aberrant functions of each individual family member. Often the effects of one Hox gene can vary due to their tissue-specific functions; the majority of Hox genes seem to lose expression during the initiation or progression of neoplasia, however there are also a significant few whose overexpression remains strongly associated with tumourigenesis. Previous observations suggest that the consequences of a given Hox gene mutation cannot be predicted a priori but will depend entirely on the nature of the mutation and on the battery of downstream genes deregulated as a result.

1.7.4. SRC-1 and HOXC11

HOXC11 has been previously associated with the differentiation of neuroblastoma GOTO cells to Schwann cells [186]. The study conducted by Zhang et al, was the first to identify the Schwann cell marker; S100β as a direct target gene of HOXC11. In terms of endocrine resistant breast cancer, mass spectrometry analysis identified HOXC11, as an interacting partner of SRC-1 in the LY2 cell line. Moreover, immunohistochemical analysis of a Tamoxifen-treated patient population showed increased localization and expression of HOXC11 in the nuclei of tumour cells. Strong associations were also observed between HOXC11 and SRC-1 in the resistant LY2 cells compared to the Tamoxifen-sensitive MCF-7 cells. This observation was again observed in the patient population. Moreover, coexpression of HOXC11 with SRC-1 was noted as a superior predictor of poor clinical

prognosis than any of the other classic parameters. This study also recognized S100β as the coactivated transcriptional target for the SRC-1/HOXC11 complex in a resistant cellular environment [62].

1.8. Hypothesis

DP McDonnell makes the point; that one of the major changes in our understanding of ERα action, is that we no longer consider the receptor itself as directly impacting the activity of the transcriptional apparatus. Rather, it serves as a nucleating point for transcriptional coregulators; proteins with different enzymatic activities that can directly modify chromatin structure and/or impact the activity of the general transcription apparatus. Thus, it is the activity of the coregulators recruited by ERα, rather than the receptor itself, that is the determinant of the response of the target gene to a specific receptor-ligand complex [128]. In particular, the SRC proteins have demonstrated the capacity to coregulate a wide and diverse range of transcription factors, both steroidal and non-steroidal and can therefore impact many transcriptional reactions across many varied biological processes. For this reason, the SRC proteins have been reclassified as master regulators. Of the 300+ coregulators which have been discovered, SRC-1 in particular has become renowned for its consistent implication with breast cancer metastasis and with therapeutic resistance. SRC-1 is a target of many signalling pathways and in response to these various pathways, has shown a clear and versatile ability to regulate unrelated genes in diverse manners to ultimately influence breast cancer progression.

Significant research has been conducted with regards to SRC-1's position in and affinity for estrogen dependent and estrogen independent signalling pathways and it is clear that its activity within these pathways is usually detrimental to the prognosis of a cancer. However, there is a significant lack of knowledge as to the downstream effectors of SRC-1 activity, the molecules and proteins which ultimately execute its tumourigenic agenda.

Global analysis of SRC-1 in conjunction with a specific SRC-1 knockdown DNA microarray has identified a host of new transcriptional targets for SRC-1 in an endocrine resistant environment. The resultant data produced a significant number of genes which were directly downregulated as a result of SRC-1 activity. Given SRC-1s staunch functional history as a coactivator, this works aims to investigate whether a master regulator such as SRC-1 has the ability to switch its functionality between coactivator and corepressor. By redefining

SRC-1 as a corepressor, this work has hypothesised that SRC-1 can have as significant an impact on tumour progression through the downregulation of "protective" genes as it could through the upregulation of its better characterised "progressive" genes.

1.9. Aims

- Using available discovery data, to investigate the genes which are potentially downregulated by the coactivator SRC-1

- To identify specific targets of SRC-1 corepressive activity and ascertain what function they may have with regards to endocrine resistance in breast cancer

- To validate the transcription factor, HOXC11 as a potential transcription factor partner for SRC-1 mediated target gene suppression

- To investigate possible mechanisms of SRC-1 mediated target gene suppression

2. Materials and Methods

2.1. Cell culture

2.1.1. Cell culture environment

Cell culture was performed in a sterile environment using a laminar airflow cabinet. All cell lines were maintained in a humid 5% (v/v) CO_2 atmosphere at 37°C. All cell culture work was carried out using standard aseptic techniques.

2.1.2. MCF-7 cell line

The MCF-7 cells were obtained from the American Type Culture Collection (ATCC) . The MCF-7 cells were originally derived from a mammary gland adenocarcinoma and retain the characteristics of a well differentiated mammary epithelium. The MCF-7 cells are an endocrine responsive cell line; they express ER and PR and are negative for HER2. The MCF-7 cells were sub-cultured in minimum essentials media (MEM) (Sigma Aldrich, Germany), which was supplemented with 10% fetal bovine serum (FBS) (Sigma Aldrich) and 2mmol/L L-glutamine (Sigma Aldrich).

2.1.3. LY2 cell line

The LY2 cell line was kindly donated by Dr. Robert Clarke, Department of Oncology, Georgetown University, DC, USA, These cells were originally developed as a stable variant of MCF-7 cells which were made resistant to LY 117018, a potent antiestrogen [187]. The LY2 cells express less ER than MCF-7 cells and have no detectable PR expression. They are also negative for HER2 expression. The LY2 cells were sub-cultured in phenol-red free MEM which was supplemented with 10% charcoal dextran stripped fetal calf serum (Sigma Aldrich), 2mmol/L L-glutamine and 10^{-8}M 4-hydroxytamoxifen (4-OHT) (Sigma Aldrich). 4-hydroxytamoxifen (4-OHT) is a metabolite of the antiestrogen, Tamoxifen. 4-OHT has a higher affinity than Tamoxifen and its other metabolites for binding to estrogen receptors and therefore, has 50 to 100-fold greater potency of inhibiting cell multiplication in breast cancer cell lines in culture [188]. The purpose of the charcoal:dextran stripping was to reduce the serum concentration of various hormones and growth factors in the LY2 media in order to make the constant exposure to 4-OHT more effective.

2.1.4. MDA-231 cell line

The MDA-MB-231 cell line was also obtained from the ATCC . The MDA-MB-231 cells were originally derived from a mammary gland adenocarcinoma. The MDA-MB-231 cells represent a typical triple negative, basal breast cancer subtype. The cells express no ER, PR

or HER2, however, they do express epidermal growth factor (EGF) as well as transforming growth factor alpha (TGF-α). The MDA-MB-231 cells are sub-cultured in Leibowitz's L-15 media (Invitrogen, Carlsbad, CA, USA) which was supplemented with 10% FBS. The MDA-MB-231 cells were incubated at 37°C in a humidified incubator. Importantly, the flasks were fitted with air tight caps in order to ensure 100% atmospheric air as opposed to 5% CO_2.

2.1.5. Routine cell culture

Breast cancer cells were cultured in T-75cm^2 filtered tissue culture flasks (Sarstedt, Germany). Cells were passaged at approximately 80% confluency by washing them twice with 5mL phosphatase buffered saline (PBS; Oxoid Limited, Basingstoke, Hampshire, England) and incubating them with 3mL of 0.05% trypsin/0.02% EDTA solution (Sigma Aldrich) for approximately 5 minutes at 37°C. The trypsin was subsequently quenched by adding 3mL of normal cell culture media to the flask. The cell suspension was transferred into a 15mL conical tube (Greiner Bio-One, Germany) and centrifuged at 60g for 4 minutes. The cell pellet was then resuspended in the required volume of normal cell culture medium and divided into new T-75 tissue culture flasks at a suitable lower confluency.

2.1.6. Cell Counting

For experiments which required a finite number of cells, manual cell counting was conducted using a haemocytometer (Neubauer, Germany). Following trypsinisation, 20µL of cell suspension was mixed with 20µL of Trypan blue exclusion dye (Sigma Aldrich) and 10µL of the suspension was pipetted onto the haemocytometer. The number of cells in the two grids on the left and on the right were averaged and multiplied by 10^4 to obtain the number of cells per mL of cell suspension. Counts were performed in duplicate and an average count per mL was calculated. The required volume of cell suspension was calculated and seeded into the appropriate cell culture vessel.

2.1.7. Endocrine treatment of *in vitro* cell lines

Prior to any experiments, involving endocrine treatments, cell lines were subcultured for 72 hours in phenol-red free MEM containing 10% charcoal dextran stripped fetal calf serum. This technique is referred to as steroid depletion and removes any potential bias which could occur from the steroid hormones present in the FBS or in the phenol red. Cell treatments were conducted at the following concentrations; 4-hydroxytamoxifen at 10^{-7}M, estradiol at 10^{-8}M and vehicle control at 0.01%.

2.1.8. Primary tumour digestion and culture

Ethical approval for all studies on human samples was obtained from the Medical Ethics Committee, Beaumont Hospital, Dublin. Biological samples were only used from patients who provided informed consent for use of their tissue in any subsequent research.

Fresh tumour samples were obtained immediately after surgical resection whereby a small sample of tumour was obtained with the consent of the pathologist. Once resected, the primary tumour was placed in a sterile petri dish inside a Tissue Culture Class II cabinet. Excess fatty tissue & blood vessels were removed from the tumour sample using sterile scalpels. The tumour sample was finely chopped into very small pieces using a sterile scalpel. The tumour was then transferred into a digestion mixture containing Dulbecco's modified Eagle's media (DMEM)/F12-Ham (Sigma Aldrich) which was supplemented with 10% FBS (Sigma Aldrich), 10% Insulin (Sigma Aldrich), 10µg/mL Penicillin /Streptomycin/ Neomycin solution (Sigma Aldrich), 1X Fungizone® Amphotericin B (Invitrogen, USA), 5µg/mL Hyaluronidases (Sigma Aldrich), 200U/mL Collagenases (Sigma Aldrich).

The tumour was incubated in the digestion mixture for at least 2 hours at 37°C on an orbital shaker set at a gentle agitation. After 2 hours, the digested tissue was left to settle in the tube for up to 5 minutes, the supernatant was then collected and centrifuged at 60g for 4 minutes. The supernatant was removed and discarded. The cell pellet was resuspended in the digestion mixture without hyaluronidase/collagenase and was centrifuged again at 60g for 4 minutes. This step was repeated once more. The cell pellet was then resuspended in primary culture media consisting of DMEM-F12 media which was then supplemented with 10 ng/mL basic fibroblast growth factor (bFGF,Sigma Aldrich), 20 ng/mL epidermal growth factor (EGF, Sigma Aldrich), 5µg/mL insulin, and 0.4% bovine serum albumin (Sigma Aldrich). The primary cultures were cultured into a non adherent T-25cm^2 filtered tissue culture flask (Sarstedt) in a 5% CO_2 humidified incubator.

2.1.9. Proliferation assay

Cultured cells were trypsinised, manually counted, and seeded at 500 cells per well in a 96-well plate (Grenier Bio One, Germany). The cells were cultured in a total volume of 250µL of normal culture media and incubated at 37°C for 24 hours. After 24 hours, the normal culture media was removed and cells were re-cultured in steroid depleted media for a further 72 hours. The cells were then treated in triplicate with vehicle (0.01% ethanol), estradiol (10^{-8}M), 4-OHT (10^{-7}M), or estradiol and 4-OHT together for 72 hours.

After the 72 hours treatment period, the cells were trypsinised again and each triplicate of treatment was pooled together. The cells were pelleted by centrifugation at 60g for 4

minutes and then manually counted. The pooled cell counts were used to determine the rate of cell proliferation of each cell line in response to endocrine stimulation.

2.2. Cellular Transfection

2.2.1. Transient siRNA knockdown of target gene expression

siRNA molecules are comprised of short double stranded RNA molecules which can cause suppression of gene activity in a sequence specific manner. The double stranded RNA is processed to 21-25bp fragments of dsRNA with dinucleotide 3' overhangs. One strand of the siRNA, called guide strand, is assembled into an RNA induced silencing complex (RISC) that cleaves the target mRNA. Target specificity to RISC is provided by the siRNA through base pairing of the guide strand with the target mRNA (Figure 2.1) [189].

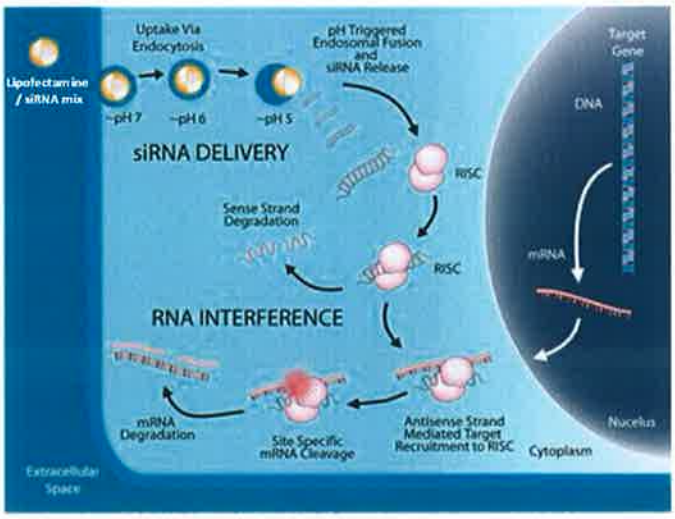


Figure 2.1. Mechanism of action for siRNA targeted silencing of gene expression in mammalian cells. *Adapted from www.cc-crs.com/maclachlan.htm*

Transient transfections of plasmid DNA or siRNA into cultured mammalian cells is carried out using a specially designed cationic lipid, known as Lipofectamine® 2000 transfection reagent (Invitrogen). The Lipofectamine® liposome molecule has a charged head group which governs the interaction between the lipid and the phosphate backbone of the nucleic acid, and facilitates DNA condensation. The liposome is also formulated with a neutral co-lipid which results in a structure with a positive surface charge when formulated in water [190]. The positive surface charge of the liposome complex mediates the interaction of the nucleic acid and the cell membrane, allowing for fusion of the liposome/nucleic acid complex with the negatively charged cell membrane. The transfection complex is thought to

enter the cell through endocytosis. Once inside the cell, the complex must escape the endosomal pathway, diffuse through the cytoplasm, and enter the nucleus for gene expression [190].

Predesigned and validated siRNA molecules were used to transiently knockdown gene expression in LY2 cells (Table 2.1). $5x10^5$ cells were seeded into a sterile 6-well tissue culture plate (Greiner Bio-One) with normal cell culture media. It is important for a successful transfection reaction that the cells are cultured in antibiotic free media for at least 24 hours before the transfection occurs. The reseeded cells were then incubated overnight at 37°C in a 5% (v/v) CO_2 incubator.

Table 2.1. siRNA molecules used for transient target gene silencing

Gene of interest	siRNA	Source	Stock concentration	Working concentration (per 6-well)
SRC-1	AM16706	Ambion	50µM	40nM
HOXC11	Hs_HOXC11_2	Qiagen	20µM	30nM
Scrambled	AM4635	Ambion	50µM	40nM/30nM

5µL of Lipofectamine® 2000 (Invitrogen) per one well of a 6-well plate was used to transfect working concentrations of siRNA and scrambled control siRNA into LY2 cells. The Lipofectamine® 2000 reagent was mixed with 245µL of Optimem (Invitrogen), and incubated at room temperature for 5 minutes. The appropriate volume of siRNA was added to a total volume of 250µL of Optimem (Invitrogen) and mixed well. The 250µL of Lipofectamine® 2000 solution was added directly to the 250µL siRNA solution and the subsequent transfection reagent /DNA mixture was gently vortexed to mix. The mixture was incubated at room temperature for 20 minutes. Following the incubation period, the media was removed from the cells in the 6-well culture plate and replaced with 1.5mLs of Optimem. The total 500µL of the transfection mixture was then added in a dropwise fashion into each well. The plate was rocked backwards and forwards and then side to side to mix the solution in the well. The cells were incubated with the transfection reagents for 5 hours at 37°C in the 5% (v/v) CO_2 incubator. Following the incubation period, the transfection/Optimem mixture was removed from the cells and 2mLs of normal culture media was added to each well. The cells were then incubated for a further 72 hours in order

to achieve a substantial knockdown of either SRC-1 or HOXC11. The loss of gene expression was validated using quantitative real time PCR.

2.2.2. Transient overexpression of target gene expression

Overexpression studies were also conducted with Lipofectamine® 2000 (Invitrogen). The overexpression transfections were transient and used plasmid vectors that had already been validated in the laboratory. The plasmid backbones used were pcDNA3.1 (Invitrogen) for SRC-1 and pc Dest47 (Invitrogen) for HOXC11 (Table 2.2).

Table 2.2. Overexpression plasmids used in transient transfections

Vector	Stock concentration	Working concentration per 6 well
pcDNA3.1 EV	0.558µg/µL	2µg
pcDNA3.1 SRC-1	2.186µg/µL	2µg
pcDest47 EV	1.14µg/µL	2µg
pcDest47 HOXC11	1.81µg/µL	2µg

EV= Empty vector

In the 24 hour period prior to transfection, 5 x10^5 MCF-7 cells were manually counted and seeded into 6-well culture plates with normal culture media. The media was kept free of antibiotics to increase the efficiency of the transfection reaction. 5µL of Lipofectamine® 2000 and 245µL of Optimem per one well of a 6-well plate was used to transfect working concentrations of plasmid vector into the MCF-7 cells. This solution was incubated at room temperature for 5 minutes. In the mean time, 2µg of the plasmid DNA was added to a total volume of 245µL of Optimem and mixed well. The 250µL of Lipofectamine® 2000 solution was added directly to the 250µL plasmid DNA solution and was gently vortexed to mix. The mixture was incubated at room temperature for 20 minutes. Following the incubation period, the media was removed from the cells in the 6-well culture plate and replaced with 1.5mLs of Optimem. The total 500µL of the transfection mixture was then added in a dropwise fashion into each well and rocked to mix. The cells were incubated with the transfection reagents for 5 hours at 37°C in the 5% (v/v) CO_2 incubator. Following the incubation period, the transfection/Optimem mixture was removed from the cells and 2mLs of normal culture media was added to each well for a total of 24 hours to ensure a sufficient overexpression of each target gene.

2.3. Gene Expression Analysis

2.3.1. RNA extraction

Cultured cells were trypsinised with 0.05% trypsin/0.02% EDTA solution, centrifuged at 60g for 4 minutes and washed in sterile PBS. The PBS supernatant was removed and the cell pellets were stored on ice.

RNA was extracted from cultured cells using the RNeasy Mini Kit (Qiagen). Appropriate amounts of RLT buffer were made up with a 1:100 β-mercaptoethanol dilution. β-mercaptoethanol was added to the RLT buffer as a preventative measure against any degrading RNases that may be released during the lysis step. β-mercaptoethanol is a reducing agent which irreversibly denatures RNases and inactivates them thus ensuring a high quality RNA yield. 350µL of RLT/ β-mercaptoethanol buffer was added to each cell pellet. The cell lysate was then passed through a 21-guage needle 5 times to homogenise the sample. An equal volume of 70% ethanol (ie 350µL) was then added to each sample and mixed with a pipette. The full volume was then added to an RNeasy spin column and centrifuged at 25,000g for 15 seconds at room temperature. The flow through was collected into a 2mL collection tube and discarded.

350µL of RW1 buffer was then added to the spin column, the tubes were spun as before and the flow through was discarded. The Rnase-free DNase kit (Qiagen) was incorporated into the protocol to ensure that all DNA was removed from the isolated RNA. 10µL of the DNase enzyme was added to 70µL of RDD buffer per sample. 80µL of the DNase mix was added onto the spin column and left to incubate for 15 minutes at room temperature. An additional 350µL of RW1 buffer was added on top of the column, it was washed through as before and the flow through was discarded.

The column was moved to a new collection tube to avoid contamination. 500µL of RPE/ethanol buffer was added to the column, the columns were spun as before and the flow through was discarded. Another 500µL of RPE/ethanol buffer was added to the column, the tubes were then spun at 25,000g for 2 minutes and the flow through was once more discarded. As a further precautionary measure against potential contamination, the empty spin column was transferred to a new collection tube and the tubes were spun once more at 25,000g for 1 minute to remove any residual buffer. The column was then placed in a 1.5mL eppendorf tube and 30µL of nuclease-free water was added to the column to elute the RNA. The column was left to stand at room temperature for 1 minute and was then spun at 25,000g for 1 minute. The RNA concentration was then determined using the Nanodrop

2000c micro volume spectrophotometer (Thermo Scientific) before the sample was stored at -80°C.

2.3.2. cDNA synthesis and reverse transcription (RT) PCR

The Superscript III First Strand Synthesis System for Reverse Transcription PCR kit (Invitrogen) was used to synthesise 1µg of cDNA from RNA. For each sample, the appropriate volume of RNA for 1µg of cDNA was added to 1µL of random hexamers, 1µL of 10mM dNTPs, and the remaining volume was made up to 10µL with nuclease-free water. Each sample was incubated at 65°C for 5 minutes and then chilled on ice for 1 minute. A second mastermix was made up (Table 2.3). 10µL of the mastermix was then added to each RNA sample and mixed well to generate a total volume of 20µL.

Table 2.3. Reverse Transcriptase master mix for cDNA synthesis.

PCR Mastermix	per sample (µL)
10X RT Master mix	2
50mM $MgCl_2$	2
0.1M DTT	2
Superscript III RT	0.5
Nuclease-free Water	3.5
RNA mix	10
Total PCR reaction	**20**

The tubes were then incubated at 25°C for 10 minutes, 50°C for 50 minutes and 85°C for 5 minutes. The PCR reaction generated a final cDNA concentration of 1µg/20µL or 50ng/µL for each sample. The cDNA samples were then stored at -20°C for quantitative PCR analysis. A RT no RNA and a no RT control were included with every cDNA synthesis performed.

2.3.3. Real time PCR analysis – Sybr Green I assay

Polymerase chain reaction (PCR) is a technique used to amplify a large amount of copies of DNA from a relatively small amount of starting DNA material. During a PCR reaction, a heat stable DNA polymerase assembles a new DNA strand from template DNA using specific DNA primers. The three main steps of a PCR are denaturation, annealing and extension, which are each repeated for a specific number of cycles.

The ABI 7500 Real Time PCR instrument (Azco Biotech, Inc, USA) was used for quantitative PCR analysis of target gene expression in the cultured mammalian cell lines. Sybr Green I technology was employed to quantify the number of amplicons produced throughout the real time PCR reaction. SYBR Green is a commonly used fluorescent agent; it emits relatively little fluorescence during DNA denaturation. However as the PCR temperature increases and polymerisation occurs, Sybr Green I intercalates the newly formed double stranded DNA amplicons and emits a strong fluorescent signal (Figure 2.2). The increase in fluorescence can therefore be directly equated to the increase in the number of DNA amplicons [191]. The Sybr Green I detection method is widely used in quantitative real time PCR analysis however the disadvantage to this method is that the emitted fluorescence is not specific to the amplified product as Sybr Green I will bind to any double stranded DNA. It is therefore important when using Sybr Green I technology to ensure accurate design and efficient use of primers to avoid issues such as primer-dimer impacting on the experimental output.

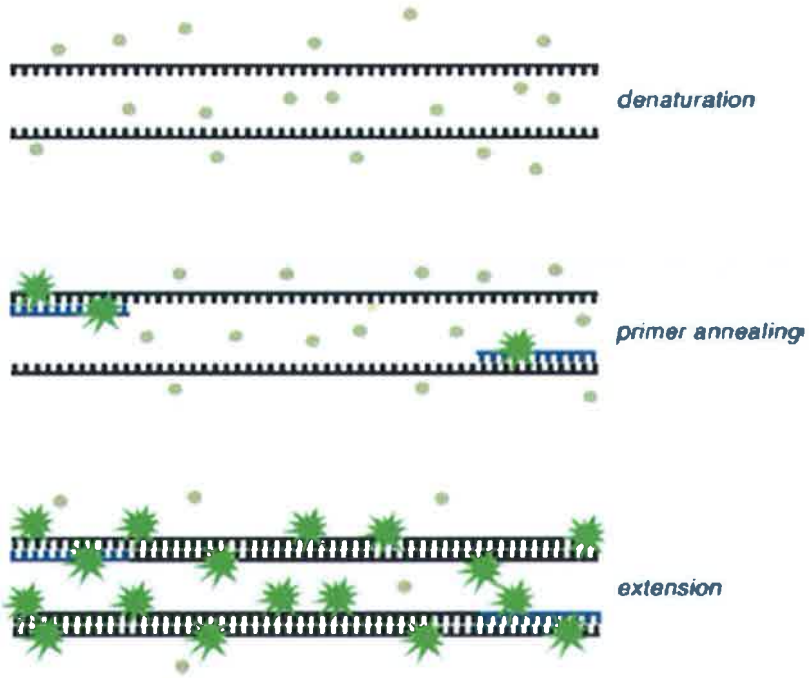


Figure 2.2. Sybr Green I mechanism of action during PCR amplification. Sybr Green I dye will only intercalate double stranded DNA. The reaction will only begin in the annealing phase and is then extrapolated during the extension phase. The fluorescence of the Sybr Green I dye increases 100- to 200-fold when bound to the minor groove of double stranded DNA. The emitted fluorescence is used to measure the amount of amplified DNA. *Adapted from Fraga et al, 2008 [191].*

Prior to setting up the Sybr Green I real time PCR experiment, the work area and pipettes were cleaned with Rnase Away Surface Decontaminant (Fischer Scientific) and bleach to further reduce any RNase activity and remove any DNA that could cause contamination of the real time run. Also all pipetting was carried out using filter pipette tips only. Experiments were set up in a Taqman 96-well gene expression plate (Applied Biosystems). 10ng of cDNA was added to each well. A Sybr Green mastermix was made up with 10µL Sybr Green enzyme (Qiagen), 1µL of 12.5µM forward/reverse primer working stock (Sigma) (Table 2.4), and made up to a total reaction volume of 20µL with nuclease-free water.

Table 2.4. Primer sequences for quantitative real time analysis.

Gene	Forward strand sequence 5'-3'	Reverse strand sequence 5'-3'	Product size (bp)
SRC-1	CATGGTCAGGCAAAAACCTT	GCTTGCCGATTTTGGTGTAT	111
HOXC11	AACACAAATCCCAGCTCGTC	AAAAACTCTCGCTCCAGTTCC	143
CD24	AACTAATGCCACCACCAAGG	CCTGTTTTTCCTTGCCACAT	188
PAWR	ACATCCCTGCCGCAGAGTGCT	TCTCGTTTCCGCTCTTTCTGCCC	70
Actin	CGGCATCGTCACCAACTG	GCCACACGCAGCTCATTG	72

Following the PCR set up, the real time amplification protocol was conducted as follows:

Table 2.5. Real time PCR protocol for Sybr Green I experiment.

Analysis mode	Cycles	Target temperature	Hold Time	Acquisition Mode
Pre incubation	1	94°C	10 minutes	none
Amplification	40	94°C	15 seconds	none
Quantification		60°C	30 seconds	none
		72°C	30 seconds	single
Melt curve	1	95°C	10 seconds	none
Dissociation		50°C	30 seconds	
		70°C		continuous
Cooling	1	40°C	30 seconds	none

A dissociating melt curve was included for all Sybr Green PCR experiments to determine the specificity of the primers for their product. A standard curve was generated for each new primer set and, all quantitative PCR runs were normalised to the β-actin reference gene and then analysed and reported using the Pfaff method of PCR quantification.

2.3.4. Real time PCR analysis – Roche Universal Probe Library (UPL) system

A different system of quantitative real time analysis was used for the *in vivo* gene expression studies. The *in vivo* work was carried out in the Xu laboratory at Baylor College of Medicine, where all quantitative PCR analysis was done with the Roche UPL system in conjunction with the StepOnePlus Real-Time PCR Instrument (Applied Biosystems).

The Roche UPL system is based on hydrolysis probe technology whereby the probe contains the fluorophore and the quencher on either end (Figure 2.3). During the annealing phase, the probe binds to the specific target sequence. Due to the proximity between the donor (fluorophore) and the acceptor (quencher) on the probe, there is no fluorescence. However,

during the extension step, the 5'-3' exonuclease activity of the polymerase hydrolyses the probe which releases the fluorophore from the quencher and the emitted fluorescence is read by the detector. During each PCR cycle, more of the released fluorescent dye accumulates, which boosts the fluorescent signal.

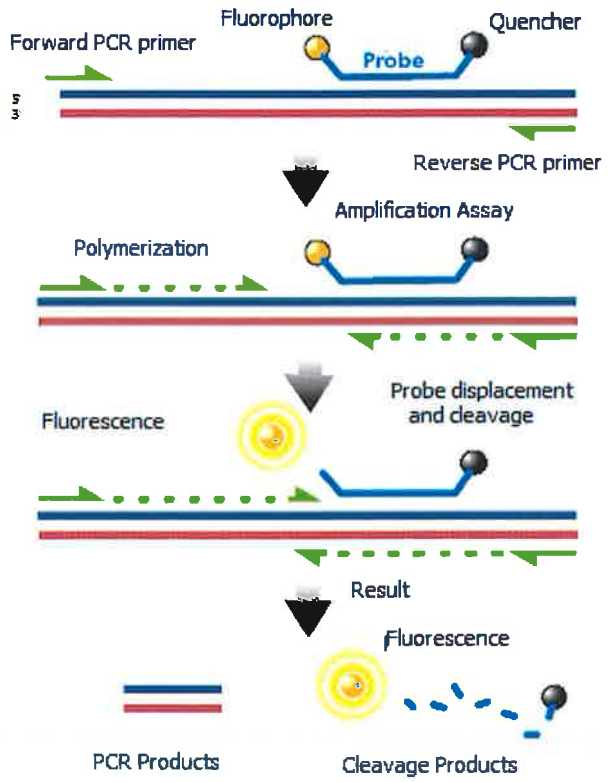


Figure 2.3. Hydrolysis probe technology for real time PCR quantification. A single probe contains two labels in close proximity to each other: a fluorescent reporter dye at the 5′-end and a (fluorescent or dark) quencher label at or near the 3′-end. When the probe is intact, the fluorescent signal is almost completely suppressed by the quenching label. When the probe is hybridized to its target sequence, it is cleaved by the 5′-3′ exonuclease activity of the DNA polymerase, which "unquenches" the fluorescent reporter dye. . *Adapted from Fraga et al, 2008 [191].*

The Roche UPL consists of 165 probes. The primers for a specific target gene are designed via the website www.universalprobelibrary.com. This website designs individual primer sets to match with a specific probe from the library. The primers can then be purchased from Sigma just as nucleotide sequences and then mixed with the designated probe before the PCR reaction takes place (Table 2.6). This system offers the same specificity as inventoried Taqman primer-probe sets but with much more flexibility and cost efficiency.

Table 2.6. Primer sequences for *in vivo* real time analysis.

Gene	Gene ID (mouse)	Forward strand sequence 5'-3'	Reverse strand sequence 5'-3'	Probe #	Size (bp)
CD24	NM_009846.2	CTTCTGGCACTGCTCCTACC	TGGTGGTAGCGTTACTTGGA	60	112
PAWR	NM_054056.2	GCAGATCGAGAAGAGGAAGC	TCTTCGTACTCATCTAAGCACTCC	7	90
18S	NR_003278.1	GCAATTATTCCCATGAACG	GGGACTTAATCAACGCAAGC	48	68

In advance of setting up the PCR 96-well plate (Applied Biosystems), the work space was cleaned and decontaminated. Again, all pipetting was carried out using filter pipette tips only. 10ng of cDNA was added to each well. A PCR mix consisting of the following reagents was made up; 1µL of 10µM forward primer working stock (Sigma), 1µL of 10µM reverse primer working stock (Sigma), 0.5µL of 10µM probe (UPL FAM-Assay) and 10µL of 2X FastStart Taqman Probe Master mix (Applied Biosystems). The total reaction volume including the cDNA was made up to 15µL with nuclease-free water. The real time PCR protocol was conducted as follows:

Table 2.7. Real time PCR protocol for Roche UPL experiments.

Analysis mode	Cycles	Target temperature	Hold Time	Acquisition Mode
Pre incubation	1	95°C	10 minutes	none
Amplification	40	95°C	10 seconds	none
Quantification		60°C	30 seconds	none
		72°C	1 second	single
Cooling	1	40°C	30 seconds	none

All of the *in vivo* quantitative PCR runs were normalised to the 18S reference gene. Analysis of the PCR run was conducted and reported using the delta delta Ct quantification method [192].

2.4. Protein Biochemistry

2.4.1. Total protein extraction

A T-75 flask of cells was trypsinised with 0.05% trypsin/0.02% EDTA and the resultant cell pellet was transferred to a 1.5mL eppendorf tube and then washed once in sterile PBS. 80µL of RIPA lysis buffer (see Appendix, Table 7.1) supplemented with a 1:100 dilution of protease inhibitor (PI; Sigma Aldrich) was added to the pellet. The pellet was resuspended and then vortexed at maximum speed for 15 seconds. The cells were then incubated on ice for 10 minutes. This was repeated twice more. The cells were then spun at 25,000g for 20 minutes at 4°C. The supernatant was removed and stored at -20°C as total protein lysate.

2.4.2. Protein quantification

A bicinchoninic acid (BCA) assay (Pierce, Il, USA) was used to assess the protein concentration in the protein lysate samples. A standard curve was prepared with a range of albumin concentrations between 0 and 1.4mg/mL as well as 1:20 dilutions of the protein samples. 25µL of each standard and protein sample was pipetted into a 96-well plate (Grenier Bio One) in duplicate. The BCA reaction solution was prepared by mixing 49 parts of reagent A (Bicinchoninic acid and tartrate in alkaline carbonate buffer) with 1 part of reagent B (4% copper sulfate pentahydrate solution). 200µL of this solution was added to each well and incubated at 37°C for 30 minutes. Absorbance was read at an optical density of 560nm on a Wallac spectrophotometer (Perkin Elmer, MA, USA). Simple linear regression was used to determine the equation of the standard curve. This equation was then applied to determine the protein content in the samples. From this concentration value, a relative volume for each protein sample was determined for Western blotting. In general, 50µg of protein was used for protein detection.

2.4.3. Western blotting

Western blotting is a technique which is commonly used to detect protein from cell lysates. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) is a technique commonly used to separate out proteins based on size. The protein lysate is separated out

on a gel and transferred to a nitrocellulose membrane by applying an electrical current. This membrane can then be stained with a specific antibody against an epitope present in the protein.

The gels were made up manually for Western blotting experiments (Table 2.8) using the ATTO electrophoresis system (ATTO, Tokyo, Japan). The gels are composed of acrylamide and bisacrylamide (Sigma Aldrich) and a cross-linking agent called N,N,N,N-tetramethyletylenediamine (TEMED; Sigma Aldrich). Gels which have a high acrylamide content have a smaller pore size and are thus used to visualise smaller proteins. A 10% resolving SDS-polyacrylamide gel was appropriate for the separation of proteins which were approximately 40kDa in size, whilst an 8% gel was generally used for proteins which were greater than 100kDa. Ammonium persulfate (APS; Sigma Aldrich) and TEMED are added to initiate polymerisation of the gel. The resolving gel was made first and was left at room temperature for approximately 30 minutes to set. The stacking gel was then made up and added on top of the resolving gel. A 1.5mm 10-well comb was inserted into the stacking gel and left for a further 30 minutes to set.

Table 2.8. Gel preparation for SDS-PAGE. Volumes cited are with regards to the preparation of 10mLs of a resolving gel and 2mL of a stacking gel for preparation of one gel on an ATTO 1.5mm plate.

	10% resolving gel (mL)	5% stacking gel (mL)
Water	4	1..4
30% acrylamide mix	3.3	0.33
1.5M Tris (pH 8.8)	2.5	-
1M Tris (pH 6.8)	-	0.25
10% SDS	0.1	0.02
10% APS	0.1	0.02
TEMED	0.004	0.002

Once the gels were set, 50µg of nuclear protein was added to the 6X loading dye and made up to a final volume of 24µL with RIPA lysis buffer. The protein samples were boiled at 95°C for 5 minutes to denature the protein and to enable a good separation of the protein on the gel. The gels were placed into an ATTO electrophoresis rig and the gel rig was filled to the

top with 1X running buffer (see Appendix, Table 7.1). 20µL of each sample was loaded onto the 10% gel along with 8µL of a Multi-Colour Broad Range Protein Ladder (Fermentas). A constant voltage of 120V was applied and the gels were run for approximately 3 hours.
Using the Bio-Rad MiniTrans wet transfer system; the proteins were then transferred to a nitrocellulose membrane for visualisation (Figure 2.4).

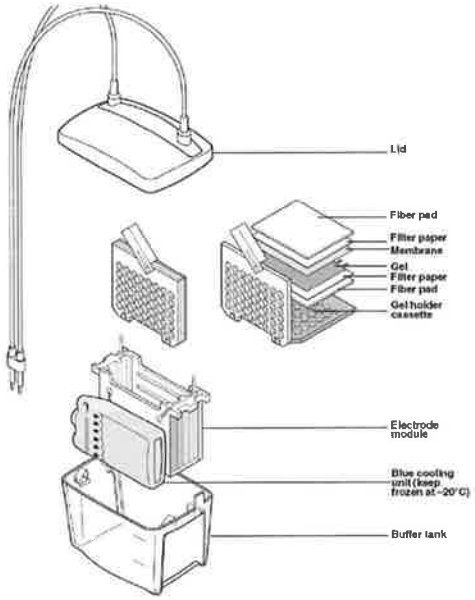


Figure 2.4. Bio-Rad MiniTrans Blot Cell Transfer System. Schematic diagram to depict the assembly of the transfer system for wet transfer of the protein samples onto the nitrocellulose membrane. *Adapted from Mini Trans-Blot Electrophoretic Transfer Cell Instruction Manual, Bio-Rad, USA*

Prior to transferring, 1X wet transfer buffer (see Appendix, Table 7.1) was made up and chilled at 4°C to improve heat dissipation throughout the transfer. The nitrocellulose membrane and four sheets of Whatman filter paper were cut to size to fit the dimensions of each gel. The fibre pads, the Whatman paper and the nitrocellulose membrane were all pre-soaked in 1X transfer buffer. The "gel sandwich" was prepared as outlined in Figure 2.4. The gel cassette was placed with the gray side down and two sheets of Whatman filter paper were placed on top of the cassette. The gel was placed on top of the filter paper, the membrane was placed on top of the gel, two additional sheets of Whatman paper were placed on top of the membrane and finally another fibre pad was placed on top of the filter paper. The cassette was closed and placed into the electrode module (Figure 2.4). The module was placed into the buffer tank and filled with ice cold transfer buffer. The transfer was then run at 90 volts (V) for 90 minutes with 1X wet transfer buffer.

Once the protein has been properly transferred to the gel, the membrane was immersed in ponceau S red solution (Sigma Aldrich) as a visual control to ensure that the protein transferred equally. The membrane was then blocked in 5% bovine serum albumin (BSA) in PBS on a roller for 1 hour at room temperature. The membrane was transferred into primary antibody (Table 2.9) and incubated overnight on a roller at 4°C.

The membrane was washed three times for 10 minutes each in 1X tris-buffered saline containing 0.1% Tween (Sigma Aldrich) (TBS-T). The membrane was transferred into horseradish peroxidase (HRP)-conjugated secondary antibody (Table 2.9). The secondary antibody was reconstituted in 5% BSA in TBS-T, and incubated on a roller for 1 hour at room temperature. Following secondary antibody incubation, the membrane was washed three times as before.

Table 2.9. Antibody concentrations for Western blotting.

Protein	Molecular Weight	Primary Antibody	Supplier	Secondary Antibody	Supplier
PAWR	40kDa	Rabbit polyclonal 1:1,000	Cell signalling 2328	Anti Rabbit IgG	Sigma Aldrich A0545
E-cadherin	120kDa	Mouse monoclonal 1:500	BD Biosciences BD610181	Anti mouse IgG	Sigma Aldrich A3682
Vimentin	57kDa	Mouse monoclonal 1:500	BD Biosciences BD550513	Anti mouse IgG	Sigma Aldrich A3682
B-Actin	42kDa	Mouse monoclonal 1:7,500	Sigma Aldrich A1978	Anti mouse IgG	Sigma Aldrich A3682

The membrane was then developed by enhanced chemiluminescence (ECL; Pierce, Il, USA). ECL substrate is used to detect HRP activity. A luminol and a peroxide solution are added in a ratio of 1:1 and the working solution is layered on top of the membrane and incubated for 1 minute. The blot was quickly placed into an X-ray cassette and exposed to X-ray film (Fuji, Tokyo, Japan) in the dark room. The film was then processed by immersion in developing solution for 2 minutes, followed by fixer solution (both Kodak, USA) for 2 minutes and then

rinsed in water. The film was left to air dry for 10 minutes and bands were visualised. The membrane was labelled and scanned.

2.5. Flow Cytometry

Flow cytometry is a widely used method for analysing expression of cell surface and intracellular molecules as well as characterising and defining different cell types in heterogeneous cell populations. Flow cytometry uses the principles of light scattering, light excitation, and emission of fluorochrome molecules to generate specific multi-parameter data from particles and cells (Figure 2.5).

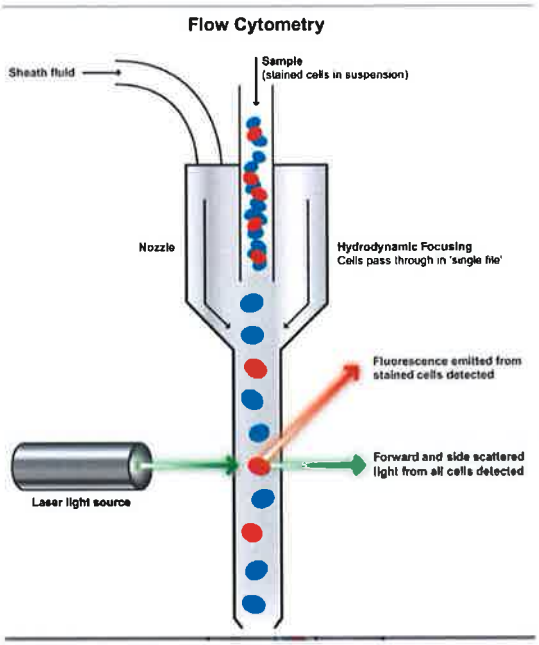


Figure 2.5. Principles of flow cytometry. Single cell solutions of stained cells are subjected to hydrodynamic focusing in the sheath fluid to ensure that the cells can interface with the laser beam. As the cells pass through the laser beam, the laser scatters light in differing directions depending on the cell type. The scattered light is then detected by a number of sensors as forward scatter (FSC) and side scatter (SSC). FSC correlates with cell size while SSC depends on the density of the cell. The laser also excites any fluorophores which are attached to the cell and measures the emitted fluorescence. *Adapted from www.cytometrysource.com*

The FACS Aria II from Becton Dickinson (BD Biosciences, San Jose, USA) is a high speed cell sorter for measuring and sorting fluorescently labelled cells. Fluorescence staining of cells is usually performed with primary antibodies that are conjugated to a fluorescent dye; the antibodies are specific to proteins of interest present on the cell surface or within the cell. The fluorescently labelled cells are hydro-dynamically focused to intercept with the laser

which can then scatter the emitted light. This light can then be detected as either forward scatter (FSC) or side scatter (SSC). FSC correlates with cell size whilst the SSC depends on the density of the cell. The combination of scattered and fluorescent light is detected and measured with a photomultiplier tube (PMT) detector for each cell; these measurements are then displayed in the form of a histogram or a scatter plot. FACSDiva software Version 6.1.2 (BD Biosciences, San Jose, USA) is installed on the FACS Aria II for operating the instrument and analysing the measurements.

2.5.1. Staining of cell lines with cell surface antibodies

Cells were trypsinised with 0.05% trypsin/0.02% EDTA (Sigma) and centrifuged at 60g for 4 minutes. The cell pellet was resuspended in normal culture media and then the cells were manually counted with a 1:1 dilution of trypan blue dye using the haemocytometer. Each sample was counted twice and an average of the two counts was taken. $5x10^5$cells/mL were counted and transferred to a sterile 1.5mL eppendorf tube. The cells were washed once with PBS and stored on ice in preparation for staining. A specific FACS wash buffer was made up with sterile PBS, 2mmol/L EDTA and 5% fetal calf serum.

Each flow cytometry experiment required certain experimental controls; an unstained sample and an isotype control for each fluorophore analysed. These controls were repeated for each cell line used.

Table 2.10. Antibodies used in flow cytometry experiments for cultured cells.

Sample	Antibody (µL)	FACS Buffer (µL)
Unstained control	-	100
PE isotype control	10	90
FITC isotype control	10	90
APC isotype control	10	90
EpCAM-APC	5	95
CD24-PE	10	90
CD24-PE &	10	80
CD49f-FITC	10	

Specific details of FACS-related antibodies are found in Appendix, Table 7.2

Both experimental controls and samples were stained with the appropriate antibody (Table 2.10). Compensation beads were also included in the experimental design in the instances where the cells were being stained for two cell surface markers to ensure that there is no bleed through of fluorescence between conjugated antibodies. The cells were incubated with the antibodies in the dark for 1 hour at 4°C. Following the incubation period, the samples and controls were spun once at 60g for 4 minutes and each cell pellet was resuspended in 500µL of FACS buffer. The samples were then transferred to 5mL Falcon round bottomed tubes (BD Biosciences) and placed on ice for analysis on the FACS Aria II.

2.5.2. FACS analysis of cell surface markers

The FACS Aria II was turned on and the Fluidics Start Up programme was run immediately. Once the instrument had completed the start-up, the 85µm nozzle was inserted and the stream was turned on and optimised in accordance with the manufacturer's recommendations. The instrument was then left for approximately an hour to allow the stream to stabilise and to allow the lasers to warm up.

BD™ Cytometer Setup and Tracking (CST) beads are run prior to each experiment as a quality control measure. The beads allow the software to automatically characterise, track, and report measurements from the FACS Aria II. Each vial of CST beads contains equal concentrations of beads of three fluorescence emission intensities: bright, mid, and dim.

The beads are used to define a baseline and run daily measurements to ensure proper working order of the cytometer. 1 drop of CST beads was added to 500µL PBS and analysed using the CST application in the FACSDiva software. Once the CST had passed without any issues or errors, the analysis could proceed.

A new experiment was then set up in the FACSDiva software and all control and sample labels were added to the new experiment. The experiment parameters were specified with regards to the number of events acquired and the flow rate. In addition, the associated graphs and statistics required were drawn. The unstained sample was loaded onto the instrument. Any adjustments that had to be made with regards to FSC and SSC were made on the unstained sample to ensure that the cells were positioned correctly for fluorescent analysis. Following the unstained sample, the PE and FITC isotype controls were loaded on the FACS Aria II and their fluorescence was measured using the appropriate channel. Gates were drawn on the fluorescence measured from these samples to define the levels of non specific staining for each fluorophore. Compensation beads were then ran on the instrument for each antibody used in the experiment to control for any overlap between fluorophores. Once the experiment was properly compensated, the CD24/CD49f stained

cells were run on the instrument and the fluorescent readings were captured for analysis. Each experiment was initially analysed in FACSDiva software but was subsequently reanalysed using a secondary software package; FlowJo (Tree Star Inc, OR, USA). Protein expression was analysed in terms of percentage expression per cell population and also in terms of mean fluorescence intensity ratio (MFIR), which reports the fluorescent intensity of a cell population relative to a specific isotype control.

2.5.3. Primary culture analysis

Primary cultures were digested as outlined in section 2.1.8. and cultured in the associated media (serum-free DMEM-F12, supplemented with 10 ng/mL basic fibroblast growth factor (bFGF), 20 ng/mL epidermal growth factor (EGF), 5µg/mL insulin, and 0.4% bovine serum albumin) for approximately 24 hours post surgery. The primary cells were centrifuged at 60g for 4 minutes and washed once with sterile PBS. The primary cells were counted manually using trypan blue dye and a haemocytometer. The cells were divided equally among the experimental samples so that each sample contained the same number of cells for analysis. The amount of antibody or isotype stain added to each sample based on the manufacturer's optimised protocol and was adjusted up or down according to the number of cells present in each sample. The cells were then stained with FITC conjugated-CD44 and PE conjugated-CD24 (BD Biosciences) for 1 hour in the dark at 4°C. Again an unstained sample and PE and FITC isotype controls were used to control the experiment as previously outlined. The experiments were set up and the cells were then analysed in the same manner as the cell lines.

2.6. Chromatin Immunoprecipitation (ChIP)

Chromatin immunoprecipitation assays determine whether a protein-DNA interaction is present at a given location within the genome. The protein-DNA complexes are fixed by crosslinking with formaldehyde and after cell lysis the chromatin is sheared, into shorter DNA fragments of approximately 500bp. Complexes containing the factor of interest are immunoprecipitated using an antibody specific to that protein. DNA is purified from the isolated chromatin, and specific genomic regions are detected using PCR (Figure 2.6).

2.6.1. Treatment of cell lines and preparation for ChIP

LY2 cells were seeded into 10cm Petri dishes (Sigma Aldrich) at a density of $1.2x10^6$cells/dish. Seven 10cm dishes were set up for each ChIP experiment, two dishes for

vehicle treated, two dishes for Tamoxifen treated, two dishes for vehicle treated IgG negative control and one dish for vehicle treated H4 positive control. The cells were subcultured in phenol-red free MEM containing 10% charcoal dextran stripped fetal calf serum and steroid depleted for 72 hours. Following the steroid depletion, two dishes were treated with 10^{-7}M 4-OHT for 45 minutes. The remaining five dishes were treated with 0.01% ethanol as a vehicle control for the same amount of time.

During the 45 minute treatment, a 1% formaldehyde solution was made up in serum free media from a 37% formaldehyde stock solution. The 1% solution was then warmed to 37°C prior to use. After 45 minutes, the media was then removed from each dish and 10mLs of 1% formaldehyde was added to each dish. The dishes were then incubated at 37°C for 10 minutes. Treatment with formaldehyde fixes the cells by crosslinking the proteins therefore any interactions between the protein of interest and the DNA as a result of the 45 minute treatment should be captured and preserved. 1M glycine (Sigma Aldrich) was added at a 1:10 dilution to the 1% formaldehyde solution on the cells to quench the fixation reaction. The dishes were incubated at room temperature for 2 minutes. The media was then poured off and the dishes were washed twice with ice cold PBS. The dishes were placed on a tray of ice and any excess PBS was removed from the cells.

1 protease inhibitor (PI) tablet (Sigma Aldrich) was resuspended in 10mLs of ice cold PBS. 1mL of PBS/PI was added to each 10cm dish and cells were scraped into a 1.5mL eppendorf tube. Each tube was centrifuged at 9,000g for 3 minutes at 4°C. The cell pellets were then stored overnight at -80°C.

2.6.2. Preparation of Dynabeads® for ChIP

Dynabeads® (Invitrogen) is a magnetic bead-based separation technology and is used in these ChIP experiments for the isolation of target proteins and the genomic regions attached to those target proteins. Dynabeads® are uniform, superparamagnetic, polystyrene beads which have either sheep anti-mouse IgG or sheep anti-rabbit IgG covalently bound to them. Using a magnetic stand (Dynal® MPC™, Magnetic Particle Concentrator; Invitrogen), to capture the Dynabeads®, the isolated target protein can be easily separated and concentrated. The protein can then be eluted off the beads for further molecular investigation.

The correct Dynabeads® were chosen depending on which antibody was being used (Table 2.11). The Dynabeads® were vortexed to ensure the solution was properly resuspended. For each sample, 50µL of Dynabeads® were added to a round bottomed 2mL eppendorf tube; this type of tube was used in order to ensure effective mixing and to minimise any

sedimentation. The tubes were placed on the magnetic stand on ice. The supernatant was removed and the Dynabeads® were washed and blocked three times with a 5mg/mL BSA in PBS solution (BSA/PBS). The Dynabeads® were then resuspended in a final volume of 350µL of BSA/PBS. The appropriate volume of antibody was then added to the 350µL. The tubes were tapped to mix and then rotated overnight at 4°C to bind the antibody to the Dynabeads®.

Table 2.11. Dynabeads® and antibodies used for ChIP experiments.

Dynabeads®	Antibody used	Antibody concentration	Dynabead volume
M-280 Sheep anti Rabbit IgG	SRC-1	6µg	50µL
M-280 Sheep anti Mouse IgG	HOXC11	12µg	50µL

2.6.3. DNA sonication and antibody incubation

The LY2 cell pellets were defrosted on ice. Three ChIP-specific lysis buffers; Lysis Buffer 1(LB1), Lysis Buffer 2 (LB2) and Lysis Buffer 3 (LB3) (see Appendix, Table 7.3) were used to lyse the cells before sonication. A PI tablet was resuspended in 10mLs of each lysis buffer prior to their incubation with the LY2 cell pellets.

1mL of LB1/PI was added to each cell pellet. The samples were then rotated at 4°C for 10 minutes. The samples were then centrifuged at 2,000g for 5 minutes at 4°C. Each cell pellet was resuspended in 1mL of LB2/PI. The samples were rotated at 4°C for 5 minutes and then centrifuged at 2,000g for 5 minutes at 4°C. Following the second centrifugation, the cells were resuspended in 600µL of LB3/PI.

Once lysed, the cells were then sonicated to shear the crosslinked DNA. The sonication protocol for LY2 cells had been previously optimised for use with the Sonifier 250 sonicator (Branson, Danbury, CT, USA) to ensure that DNA was sheared into pieces that were between 200-1000 base pairs in length. The sonicator settings used for an LY2 ChIP was output 4 and duty cycle set at 90%. The LY2 cells were subjected to a total of eight, 10 second pulses for each sample with a 1 minute interval on ice between pulses to allow dissipation of any heat generated. Following sonication, 60µL of 10% Triton X-100 (Sigma Aldrich) was added to the sonicated lysates and then centrifuged at 20,000g for 10 minutes at 4°C. The DNA was quantified on the Nanodrop 2000c spectrophotometer and the samples were then adjusted so that each contained 200µg of DNA. The samples were then diluted to a final volume of

1mL with LB3/PI and a final concentration of 1% Triton X-100. 50µL of each sample was taken and stored at -80°C as "Input" controls.

The antibody/Dynabeads® complex from the previous day, were captured using the magnetic stand and the supernatant was discarded. The beads were washed three times in BSA/PBS to remove any unbound antibody and resuspended in a final volume of 200µL of BSA/PBS. The 200µL of beads was then added to each sonicated ChIP DNA sample and rotated overnight at 4°C.

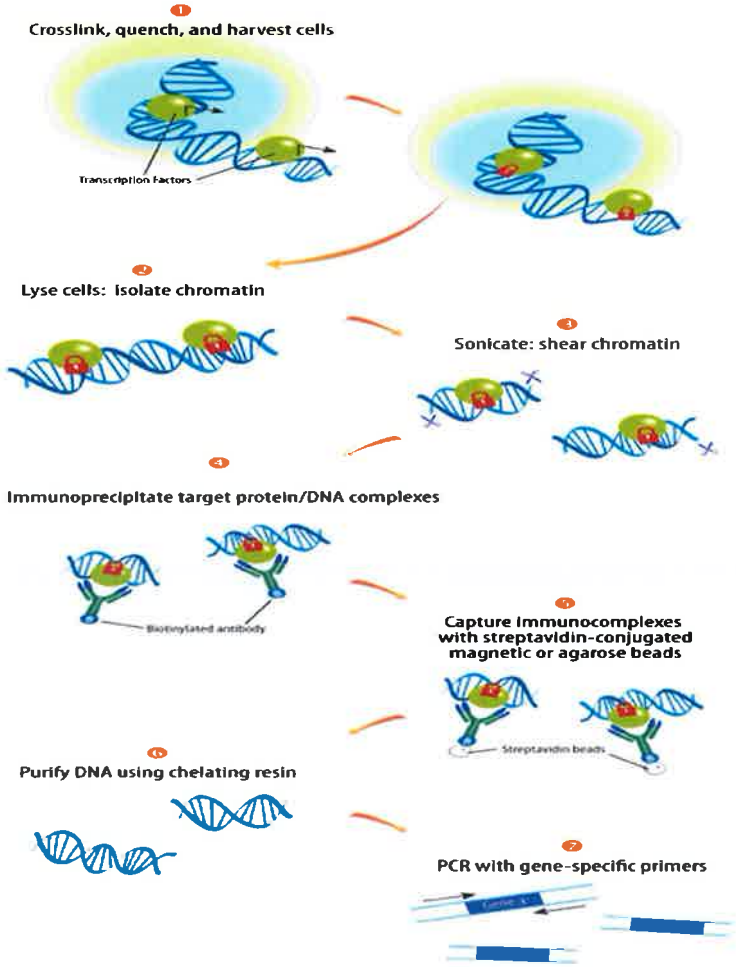


Figure 2.6. Principle of chromatin immunoprecipitation reaction. *Adapted from www.rndsystems.com*

2.6.4. Reverse crosslinking of the immunoprecipitated DNA

A 50mL quantity of radioimmunoprecipitation (RIPA) wash buffer was made up immediately prior to use (see Appendix, Table 7.3). The samples were removed from their overnight incubation in the cold room and the beads were captured using the magnetic stand. The

beads were resuspended in 1mL of RIPA buffer and gently mixed. The beads were then captured with the magnetic stand and the supernatant was removed. This washing step was repeated 8 times. Once the final supernatant was discarded, the beads were resuspended in 1ml of 1X TE buffer (see Appendix, Table 7.3) and were gently mixed to wash. The beads were captured once more with the magnetic stand and the TE supernatant was removed. The samples were then centrifuged at 800g for 3 minutes. Any excess supernatant was removed from the pelleted beads with a pipette. 200µL of elution buffer (see Appendix, Table 7.3) was added to the immunoprecipitated DNA. The Input samples were removed from the -80°C freezer and 200µL of elution buffer was added to these samples as well. All of the samples were then placed in a 65°C water bath overnight for a maximum of 16 hours to reverse cross link the DNA.

2.6.5. DNA purification

The samples were removed from the water bath, they were briefly vortexed and quickly centrifuged. 150µL of Proteinase K mixture (see Appendix, Table 7.3) was added to each sample and gently mixed. The samples were then incubated for 2 hours at 37°C.

300µL of phenol chloroform:isoamyl alcohol (24:1) (Sigma Aldrich) was added to each sample to extract the DNA. The samples were mixed briefly and centrifuged at 12,000g for 5 minutes at room temperature. The upper aqueous phase was removed and transferred to a new 1.5mL eppendorf tube. Twice the volume (~700µL) of ice cold 100% molecular grade ethanol was added to each sample. The tubes were vortexed to mix and then incubated for 30 minutes at -80°C. All samples were then centrifuged at 25,000g for 20 minutes at 4°C.

The supernatant was removed and the DNA pellet was washed with 500µL of 70% ethanol. The samples were centrifuged at 25,000g for 5 minutes at 4°C. The supernatant was removed and the DNA pellet was dried by centrifugation in a SpeedVac vacuum centrifuge (Eppendorf) at 30°C for 5 minutes. Finally, the DNA pellet was resuspended in 30µL of nuclease-free water and stored at -20°C.

2.6.6. Real time quantification of isolated DNA

The amounts of immunoprecipitated DNA was quantified using the Sybr Green I assay and the ABI 7500 Real Time PCR instrument. Primers were designed against specific regions of the promoter for target genes of interest (Table 2.12).

Table 2.12. Primers for quantification of immunoprecipitated DNA with real time PCR.

Gene	Forward strand sequence 5'-3'	Reverse strand sequence 5'-3'	Product size (bp)
CD24	GCGTGAGTTATTATTGGCTAAGGT	GATCACATGGTCAGGAGATCG	264
PAWR	CTAGCAGCTTCGCGGCGTCC	GAAGGGCGTGGAGTGCCGTT	71

1µL of ChIP DNA or Input was added to a 96-well Taqman PCR plate. A mastermix composed of 10µL of Sybr Green I, 1µL of a 12.5µM forward/reverse primer working stock (Sigma Aldrich) and 8µL of nuclease-free water was added to each well. The Taqman plate was briefly centrifuged to mix the contents of each well. The plate was then placed on the PCR machine and the PCR protocol from Table 2.5 was used to quantify the DNA expression. The resultant Ct values were calculated relative to the IgG control, and the data was then normalised to specific experimental replicates.

2.7 Immunohistochemistry

Immunohistochemistry (IHC) is used to demonstrate the presence and location of proteins in tissue sections. It enables the observation of proteins in the context of intact tissue. Immunohistochemistry was performed on previously constructed tissue micro-arrays (TMA) that consisted of 0.8mm diameter cores of patient tumour specimens.

The TMA slides were deparaffinised by passage through two xylene baths for a total of 3 minutes each followed by rehydration by sequential passage for 3 minutes each through decreasing concentrations of industrial methylated spirits (100% → 90% → 70%). The slides were then washed twice in PBS for 5 minutes. Endogenous peroxidase activity was quenched by incubating the slides in 3% hydrogen peroxide (H_2O_2, Sigma Aldrich) for 20 minutes followed by one wash in PBS.

Heat mediated antigen retrieval was performed by placing the slides in a 10mM sodium citrate buffer (pH 6.0) and heating in a closed plastic container in a domestic microwave for 7 minutes at high power and then 5 minutes at medium power followed by 10 minutes cooling at room temperature. The slides were then blocked in a 10% solution of goat serum made up in PBS for 90 minutes at room temperature. Blocking the slides ensures that any non specific binding of secondary antibody to endogenous immunoglobulins is kept to a

minimum. Following blocking, the slides were incubated with the primary antibody, individual concentrations and incubation times were determined according to the manufacturer's instructions and subsequent optimisation (Table 2.13). The primary antibodies were all diluted in PBS to the required concentration.

Table 2.13. Primary antibody condition used for immunohistochemistry of TMA slides.

Target Protein	Blocking Solution	Primary antibody	Primary antibody concentration	Primary antibody incubation time	Secondary Antibody
SRC-1	Goat serum	Rabbit polyclonal anti SRC-1	1 in 100 (2μg/mL)	2 hours at room temperature	Anti rabbit IgG
HOXc11	Goat serum	Chicken polyclonal anti HOXc11	1 in 500 (2μg/mL)	90 minutes at room temperature	Anti chicken IgG
CD24	Goat serum	Mouse monoclonal	1:50 (1mg/ml)	2 hours at room temperature	Anti mouse IgG
PAWR	Goat serum	Rabbit polyclonal	1:25 100μg/mL	Overnight at 4°C	Anti rabbit IgG

Following primary incubation, the TMA slides were washed in PBS containing 0.1% Tween (PBST, Sigma Aldrich) on a rocker for 5 minutes. This was repeated for a total of three washes. The slides were then incubated in the relevant Vectastain Elite (Burlingame, CA) biotinylated secondary antibody (PK-1600 series) at a 1:200 dilution in a solution of PBS containing 10% of the goat blocking serum for 30 minutes at room temperature. The signal was then amplified by incubating with the Avidin-biotin complex from the Vectastain Elite kit for a further 30 minutes. The staining was developed using 3,3'-Diaminobenzidine tetrahydrochloride (DAB) (Sigma-Aldrich) on the slides for 3 minutes followed by a 3 minute incubation with haemotoxylin (Sigma-Aldrich) for nuclear counterstaining. The slides were then washed in flowing tap water for 5 minutes. The slides were dehydrated again by passing through increasing concentrations of IMS (70% → 90% → 100%) for a 3 minute period per bath and into two xylene baths for another 3 minutes in each. Finally, the slides were mounted with DPX mounting solution (Sigma-Aldrich) and left to dry for 10 minutes before storage at room temperature.

The stained TMA slides were called by two observers using light microscopy and scored using the Allred scoring system [193]. This scoring system assigns two scores: a proportion score to represent the area of tissue stained (none=0, <1%=1, >1%<10%=2, >10%<33%=3,

>33%<66%=4, >66%=5) and an intensity score that represents the average intensity of the positive tumour cells (none=0, weak=1, intermediate=2, strong=3). The scores area summed to obtain a total score between 0 and 8. A total score of greater than 3 was deemed positive. IHC images were captured using an Olympus IX51 light microscope and Cell^F imaging software (Olympus).

2.8. DNA Methylation Array

DNA methylation is one of the most commonly occurring epigenetic events in the mammalian genome. DNA methylation refers to a covalent chemical modification, resulting in the addition of a methyl (CH_3) group at the carbon 5 position of the cytosine pyridine ring. DNA methylation affects only cytosine residues and is specific for CpG sequences [194]. CpG sequences are regions of DNA where a cytosine nucleotide occurs next to a guanine nucleotide in the linear sequence of bases along its length. The "p" refers to the one phosphate link between the two nucleotides. There are regions of the genome that have a higher concentration of CpG sites, known as CpG islands. Methylation of CpG islands which are located within the gene promoter region has been associated with gene silencing and has thus been proposed as an important mechanism of transcriptional regulation [195]. In order to assess the ability of SRC-1 to use methylation as a means to transcriptionally regulate its downstream target genes, genome-wide methylation levels were assessed in LY2 cells which had been transfected with either scrambled control or with siRNA against SRC-1. In order to do this, methylated DNA immunoprecipitation was coupled with DNA methylation microarrays (MeDIP-ChIP).

2.8.1. Cell line preparation and extraction of DNA for MeDIP

A transient knockdown of SRC-1 was carried out in the LY2 cells as previously described in Section 2.2.1. Following the 72 hour transfection, the LY2 cells were trypsinised with 0.05% trypsin/0.02% EDTA solution and centrifuged at 300g for 5 minutes. The cells were washed once in sterile PBS, transferred to a 1.5mL eppendorf tube and then resuspended in 200µL of sterile PBS for DNA extraction with the DNeasy Blood and Tissue Kit (Qiagen). 20µL of proteinase K was added to each sample. 200µL of Buffer AL was also added to each sample, the tubes were vortexed immediately to mix the contents. The samples were incubated at 56°C for 10 minutes. 200µL of 100% ethanol was added to each sample and mixed again. The mixture was transferred into a DNeasy Mini spin column which was placed into a 2mL collection tube. The columns were centrifuged at 6,000g for 1 minute. The flow-through and

collection tube were both discarded. The spin column was placed into a new 2mL collection tube and 500µL of Buffer AW1 was added. The columns were centrifuged again at 6,000g for 1 minute. Again both the flow-through and the collection tube were discarded. The column was placed into a new 2mL collection tube and 500µL of Buffer AW2 was added to each. The samples were then centrifuged at 20,000g for 3 minutes to dry the DNeasy membrane. The flow-through and the collection tube were both discarded. The spin column was transferred to a 1.5mL eppendorf. 200µL of Buffer AE was added to the column and left to incubate at room temperature for 1 minute. The spin column was centrifuged at 6,000g to elute the DNA. The elution step was repeated with the same 200µL to increase the DNA yield. The extracted DNA was stored at -20°C for the MeDIP experiment.

2.8.2. Sonication of extracted DNA for MeDIP

The MeDIP reaction was set up and carried out as first described by Weber, *et al.* [196]. The first step prior to immunoprecipitation was sonication of the DNA being used. The DNA should be fragmented into pieces that are between 200-800bp in length in order for a successful MeDIP reaction to take place. 5µg of transiently transfected LY2 DNA was sonicated with Branson S-250D sonicator using the optimal settings of 50% amplitude with 8 pulses of 10 seconds per cycle. Each sample was mixed well and received 5 sonication cycles with a gap of 1 minute between each cycle.

Prior to sonication, the tip of the sonicator was wiped thoroughly using a lint free tissue, in order to minimize any cross contamination. An eppendorf tube containing nuclease-free water was applied to the sonicator tip as an experimental blank. Between different samples, the blanking step using nuclease free water was carried out as aforementioned.

Once sonicated, the DNA was then visualized by resolving on an agarose gel. This was carried out using the E-gel™ system (Qiagen, G5018-02). Samples were prepared for loading, by mixing 5µL of the sonicated DNA with 1µL of gel loading dye (0.25% bromophenol blue, 0.25% xylene cyanol FF and 30% glycerol in water). The samples were loaded onto a pre-cast 2% agarose E-gel containing ethidium bromide (Qiagen). The comb was carefully removed from the gel and the samples were loaded into each well of the gel alongside a 123bp ladder (Invitrogen, 15613011), leaving the first and the last well empty. The empty wells were filled with nuclease free water. The gel was then run at a pre-set setting for the 2% gel on this system for 35 minutes. The sonicated DNA was visualized using a gel documentation system (AlphaImager Mini, 92-14005-00).

2.8.3. Immunoprecipitation reaction and elution of enriched fragments

From each sonicated LY2 sample, 500ng of DNA was kept aside as the input DNA. The remaining sonicated fraction was then used for the immunoprecipitation reaction. The MeDIP fraction was first diluted in 450µL 1X TE buffer (see Appendix, Table 7.3) in a 2mL sterile round bottom eppendorf. The DNA was then denatured by incubating at 98°C for 10 minutes followed by an immediate incubation on ice for 10 minutes. 51µL of 10X IP buffer (100mM Na-Phosphate pH 7.0, 1.4 M NaCl, 0.5% Triton X-100; 1M Na Phosphaste: 39mL 2M NaH_2PO_4, 61mL 2M Na2HPO4 and 100mL water) was added to the tube, followed by the addition of 10µL of 5'-methylcytosine antibody (Eurogentec #BI-MECY-1000). The eppendorf tubes were sealed using Parafilm and incubated for 17 hours at 4°C on an orbital rotor, at a 70° inclination angle to allow effective mixing of the antibody and buffer with the sonicated DNA.

Following this incubation period, an eppendorf tube with 40µL of sheep anti-mouse Dynabeads® (Invitrogen) were prepared for each sample. The beads were pre-washed using 800µL 0.1% BSA-PBS for 5 minutes at room temperature, on a bench top thermomixer (Eppendorf) at maximum speed. The beads were then captured using the magnetic stand. The pre-wash step was repeated once more to ensure complete removal of any storage buffer residues. The beads were then added to each MeDIP sample tube containing the 5'-methylcytosine antibody and the sonicated DNA. The mixed samples were then incubated at 4°C for 2 hours using the same rotor settings as previously described. After this incubation period, the beads were captured using the magnetic stand, the supernatant was removed and the beads were washed three times using 700µL of 1X IP buffer for 10 minutes at room temperature.

To ensure complete removal of any unbound fraction from the tubes, the bead suspension was mixed thoroughly by pipetting a minimum of 10 times. The Dynabeads® were captured for a final time with the magnetic stand and any residual IP buffer was removed. The samples were resuspended with 250µL of proteinase K digestion buffer (50mM Tris pH 8.0, 10mM EDTA, 0.5 % SDS) and 7µL of 10mg/mL proteinase K was added to each tube. The samples were incubated at 37°C for 3 hours using a thermomixer (Eppendorf) with constant agitation to prevent sedimentation of the beads. Following this step, 250µl of TE saturated phenol solution (Sigma Aldrich) was added to the tubes and then vortexed. The tubes were then centrifuged at 20,000g for 5 minutes. The upper aqueous phase was removed and added to an empty eppendorf tube. To this tube, 250µL of chloroform:isoamyl alcohol (24:1) (Sigma Aldrich) was added and vortexed, followed by centrifugation at 20,000g for 5

minutes. After the centrifugation, the upper phase was removed and collected into a new eppendorf tube. The DNA was precipitated out of the solution with the addition of 20µL of 5M NaCl, 1µL glycogen and 2X volumes of 100% ethanol (~ 500µL), and the tube was vortexed thoroughly. The tubes were centrifuged at 20,000g for 10 minutes. The supernatant was carefully removed to ensure the DNA pellet remained intact. The DNA pellet was re-suspended in 200µL of ice cold 80% ethanol; the solution was mixed thoroughly and centrifuged at 20,000g for 10 minutes. After centrifugation, the supernatant was removed carefully and the DNA pellet was dried by centrifugation in a SpeedVac vacuum centrifuge (Eppendorf) at 30°C heat for 5 minutes. Finally, the pellet was re-suspended in 60µL 1X TE and subsequently stored at -20°C until further use.

2.8.4. CpG island-promoter plus microarray

Genome-wide methylation assessment was carried out using a CpG island promoter plus array (Human Meth 385K Prom Plus CpG, Roche NimbleGen). The 385,000 probe CpG Island Plus Promoter Array is a single array design which covers all UCSC (HG18) annotated CpG islands and promoter regions for all RefSeq genes. Promoter regions are covered by 1Kb of tiled sequence, while small CpG islands are extended at both 5' and 3' ends for a total coverage of 700bp for more reliable detection. DNA methylation positive control regions, such as the *HoxA* gene cluster, *H19/IGF2, KCNQ1* cluster and *IGF2R* gene locus are also included on the array.

2.8.5. Labelling of DNA with Cy Fluorophores

Following the MeDIP reactions, the input and the immunoprecipitated DNA were quantified using the Nanodrop ND-1000 spectrophotometer (Thermo Scientific). These fractions were differentially labelled using a Dual-Color DNA Labeling Kit (Roche NimbleGen). The input DNA was labelled using Cy3 and the immunoprecipitated DNA fraction was labelled using Cy5, both were set up in 1.5mL thin-walled PCR eppendorf tubes (Table 2.14).

Table 2.14. Components of MeDIP labelling reactions.

Reagents	MeDIP Samples	Input Samples
DNA Concentration	200ng	200ng
Diluted Cy3-Random Primers	N/A	40µL
Diluted Cy5-Random Primers	40µL	N/A
Nuclease-free Water	Adjust to 80µl	Adjust to 80µl

The labelled samples were theat denatured at 98°C for 10 minutes before chilling on ice for 2 minutes.

The dNTP master mix was prepared as outlined in Table 2.15 and 20µl of this master mix was added to each of the labelled samples. The reaction was mixed thoroughly with a pipette and incubated at 37°C overnight on a heat block. The heat block was covered with tin foil to avoid degradation of the fluorophores via light exposure.

Table 2.15. Components of dNTP/Klenow mastermix.

Reagents for Master Mix	Volume per sample
10mM dNTP Mix	10µL
Nuclease-free Water	8µL
Keno Fragment (3'>5' exon-) 50U/µL	2µL

The reaction was stopped with addition of 10µl of 0.5M EDTA, followed by 11.5µl of 5M NaCl and 110µl of 100% isopropanol to each sample. The samples were vortexed after addition of each of the components to aid complete mixture of the reaction mixture and effective precipitation. The samples were incubated for 10 minutes at room temperature and protected from light. The samples were then centrifuged at 12,000g for 10 minutes. The supernatant was removed using a pipette. The appearance of the pellet was either pink (Cy3) or blue (Cy5) depending on the dye. The pellet was rinsed with 500µL of 80% ice-cold ethanol. The pellet was dislodged by gently flicking the tube. The tube was centrifuged at 12,000g for an additional 2 minutes. The supernatant was removed again by gently inverting the tube on tissue and dried on a SpeedVac vacuum centrifuge (Eppendorf) at 30°C for 10 minutes, protected from light.

Following centrifugation, 25µL of nuclease-free water was added to rehydrate the pellet for 5-10 minutes in the dark. The tube was gently flicked to ensure complete rehydration of the pellet. The samples were then vortexed and centrifuged at low speed. Each sample was quantified again using the Nanodrop ND-1000 spectrophotometer (Eppendorf). 6µg of the labelled MeDIP DNA and 6µg of the labelled Input DNA were combined in a clean 1.5mL eppendorf. The remaining labelled DNA was dried in a vacuum centrifuge at 30°C for 15 minutes and stored at -20°C protected from the light, until needed.

2.8.6. Loading and Hybridization of labelled MeDIP DNA

Following the labelling reaction, the samples were prepared using the Hybridization Kit (Roche NimbleGen). The MAUI (Micro Array User Interface) hybridisation system was used to hybridise the labelled DNA to the microarrays. The system was switched on and set to 42°C and allowed to stabilise for 3 hours before use. The combined MeDIP/Input sample was then concentrated down to a final volume of 5µl using the vacuum centrifuge. The hybridisation mastermix was made up (Table 2.16) and 13µl was added to each 5µL sample. The samples were mixed by vortexing, followed by a brief centrifugation to collect the sample at the bottom of the tube.

Table 2.16. Components of the hybridisation master mix.

Components	385K array
2X Hybridisation buffer	11.8µL
Hybridisation Component A	4.7µL
Alignment Oligo	0.5µL

The pellets were next incubated at 95°C for 5 minutes, protected from light, and then transferred to the MAUI where they were left to incubate at 42°C for at least 5 minutes or until ready to use.

The NimbleGen Hybridization System requires NimbleGen mixers, which adhere directly to the microarray slides before sample loading and hybridization (Figure 2.7). The mixers allow for low hybridization volume, which ensure a greater sensitivity and more reliable results. They fit NimbleGen arrays precisely therefore for the 385K array, the X1 mixer (Roche NimbleGen) was used.

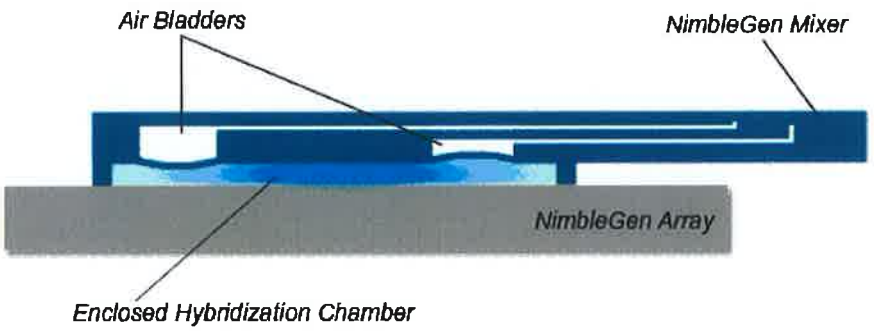


Figure 2.7. Schematic diagram of the NimbleGen hybridisation system. *Adapted from www.nimblegen.com*

The X1 mixer was removed from the packaging, and placed on the Precision Mixer Alignment Tool (PMAT) according to the manufacturer's instructions. The glass array slide was carefully removed from the dessicator and carefully placed on the PMAT in the correct orientation to allow the application of mixer. Using a pair of tweezer, the plastic backing of the adhesive gasket was removed from the mixer and the PMAT was closed in order to allow the gasket to make contact with the array slide. The mixer was gently pressed onto the slide to allow complete adhesion of the mixer to the slide. The mixer-slide assembly was removed from the PMAT and a mixer brayer was used to seal the mixer carefully avoiding any bubbles or loose ends in the assembly. The slide was then placed into the hybridisation system and incubated at 42°C for 5 minutes until the sample was loaded.

All samples were loaded onto the array using a Gilson Microman positive displacement pipette. 16µl of sample was carefully loaded onto arrays. The sample was added to the fill port in a very gradual manner, avoiding any air bubbles in the array. Once filled completely, any overflow from the fill or vent ports were carefully removed using a lint free tissue. Adhesive port seals were applied to the fill and vent port using a tweezer. The seal was gently pressed. The clamp of the hybridisation chamber was closed gently and the mixing system was turned on. Samples were allowed to hybridise to the arrays at 42°C for 16-20 hours.

2.8.7. Post hybridisation microarray washes

Following hybridisation, the arrays were washed and processed using the Wash Buffer Kit (Roche NimbleGen). The wash buffers were prepared as outlined in Table 2.17, prior to removing the arrays from the hybridisation system.

Table 2.17. Preparation of wash buffers.

Reagents	Wash I for Bath	Wash Tank I	Wash Tank II	Wash Tank III
Ultrapure water	225mLs	225mLs	225mLs	225mLs
10x wash buffer I	25mLs	25mLs	-	-
10x wash buffer II	-	-	25mLs	-
10x wash buffer III	-	-	-	25mLs
1M DTT	25µL	25µL	25µL	25µL

A shallow water bath was prepared with 250mLs of wash buffer I pre-heated to 42°C. The other washes were prepared in the wash tanks provided with the kit and kept at room

temperature. Following this, the slide was removed from the MAUI and loaded onto the mixer disassemble tool, which was then placed into the pre-warmed wash buffer bath. Once submerged completely, the tool along with the slide were gently agitated in the bath, while gently peeling the mixer from the tool and then discarded. The slide was washed for an additional 2-3 minutes. After this, the slide was removed from the tool and any excess of the wash solution was removed by gently dabbing with a lint free tissue.

The slide was then placed in wash tank I and the lid of the tank was closed. The slide was then washed by vigorous and constant agitation of the wash tank for 2 minutes. Once washed, the slide was removed from the wash tank using a blunt end forceps, followed by blotting the edges on a lint free tissue. The slide was transferred to wash tank II and washed in the same manner as wash tank I for 1 minute. This was followed by a final wash in wash tank III for 15 seconds. All slides were processed using these steps, processing 3 slides at a time. The slides were washed thoroughly in order to eliminate any residual sample or Cy fluorophores, thus minimising the background. Once the slide was washed using wash buffer III, the slide was blotted on a lint free tissue. The slide was then placed on a mini-centrifuge which was fitted with an array slide attachment. The slide was spun dry for 1-2 minutes. The arrays were then stored in their original plastic covers prior to scanning.

2.8.8. Scanning

The washed and dried arrays were scanned using the Axon 4000B scanner, and the images generated were subsequently acquired using GenePix Pro 6.0 software. The microarray slide was placed into the slide carriage in the correct orientation. The scanner software was initiated and settings adjusted as outlined in Table 2.18.

Table 2.18. PMT scan settings for two-colour scanning.

Settings	Adjustments
Cy3 Wavelength	532nm
Cy5 Wavelength	635nm
532 PMT Gain	650
635 PMT Gain	750
Power (%)	100%
Pixel size (μM)	5μM
Lines to average	1

In the "Image" option on the left side of the software screen, "Ratio" was selected to allow simultaneous scanning of both channels: Cy3 and Cy5. A preview scan was performed to locate the array on the glass slide, also allowing adjustment of the coordinates to set out the area to be scanned. In order to obtain a balanced dual-colour image, the photomultiplier tube (PMT) settings were adjusted for both channels based on the paired fluorescence intensity for both channels. While adjusting the PMT settings, utmost care was taken to avoid saturation and reduce noise. The images generated for both the channels were saved as high resolution TIFF images. After scanning and data acquisition, the arrays were stored in a dessicator for further scanning if required.

2.8.9. Nimblescan Data analysis

Following data acquisition, microarray analysis was performed using the nimblescan software version 2.3.78 according to the manufacturer's instructions for each of the processed array formats. The Nimblescan software provided by NimbleGen analysed the TIFF images and assigned a numerical value to the signal intensities of the Cy3 and Cy5 channels for every feature on the array. The values for Cy5 (immunoprecipitated DNA) were then divided by Cy3 (input fraction) in order to obtain a ratio file for each feature. These files were then processed to generate final methylation peak, which were used for subsequent application and analysis.

2.8.10. MeDIP-ChIP analysis

The signal intensities for all probes on the array were saved as pair files (.txt), which was the raw data format for the DNA methylation experiment. Each feature on the array has a corresponding scaled log2 ratio, which is the ratio of the input signals from Cy5 and Cy3 samples that are co-hybridized to the array. The log2 ratio was then computed and scaled to centre the ratio data around zero. This centering was performed by subtracting the bi-weight mean for the log2 ratio values for all features on the array from each log2-ratio data. From these scaled log2-ratio data, a fixed-length window, set at a default of 750bp was placed around each consecutive probe. Next, a one sided Kolmogorov-Smirnov (KS) test was applied to determine whether the probes were drawn from a significantly more positive distribution of intensity log-ratios than those on the rest of the array. The resulting score for each probe is the –log10 p-value from the windowed KS test around that probe. These p-value files were then used to generate final peak files. The NimbleScan software detects peaks by searching for the probes specified by the user, and above a minimum p-value cutoff of –log10 p-value = 2. These peaks were then merged if they were located within

500bp from each other. This resulted in the generation of a final peak file. These peak files along with the KS test (p-value data) and the raw ratio files were visualized on SignalMap software.

2.9. Bioinformatic analysis

Bioinformatic analysis was performed with the assistance of our colleagues, Yuan Hao and Dr. P. O'Gaora, in the Department of Bioinformatics, Conway Institute, UCD, Dublin 4. SRC-1 chromatin immunoprecipitated DNA samples were sequenced on Illumina Genome Analyzer II. Short reads of 36bp that passed the default Eland quality filter were aligned to human genome (GRCh37/hg19) with Bowtie (version 0.12.5) requiring singleton in the top stratum and up to 2 mismatches [197]. Peaks were called using MACS program (version 1.4.1) with optimized shifting size to merge the enriched regions on two opposite strands to form the final peak profile [198]. Uneven numbers of reads were balanced by scaling input to chipped sample during peak calling. Heatmaps of binding events of Tamoxifen treated samples with corresponding regions from LY2 and MCF-7 [199] vehicle treated samples were generated as described in Ross-Innes et al, 2010 [200]. Peak annotations were conducted using the ChIPpeakAnno package available in R/Bioconductor (version 2.9/2.14) [201]. Peaks overlapping with genes were further assessed using PeakAnalyzer to show the precise binding position on gene structure [202]. Binding site distribution around the transcriptional start site (TSS)/ transcriptional termination site (TTS) were plotted using the CEAS program [203]. Protein coding and non-coding genes were defined based on Ensembl genome annotation (Ensembl Genes 66, GRCh37) with R scripts. HOXC11 motif was adopted from the UniPROBE database [204]. The first 4bp and the last 3bp of this motif were deemed to have low-information content and removed from further consideration, resulting in a motif of 8bp being employed for binding site discovery. Python scripts and TFfinder were used to perform motif instance searching in promoter regions (upstream 1000bp) [205]. The Affymetrix DNA microarray data was evaluated with R/Bioconductor packages. Differentially expressed genes were further determined as to whether or not they have ChIPseq peak(s) in the genomic region 5000bp upstream of the transcriptional start site and/or within the first exon. ChIPseq peaks in the upstream region and/or in the first exon were determined using BEDTools [206]. Pathway analysis was conducted using Metacore software (GeneGo Inc, CA, USA).

2.10. Statistical analysis

Anonymous databases were maintained on Microsoft Excel, (Microsoft, Redmond, WA, USA). Multivariate analysis was performed using STATA 12 data analysis software (Stata Corp. LP, Texas, USA) and Fischer's exact test for two by two tables was used for comparison of categorical data. Survival times between groups were compared using the Wilcoxon test adjusted for censored values. Statistical analyses were carried out using GraphPad Prism 5 software (La Jolla, CA, USA). P values <0.05 were considered significant. Two sample Student's t test or One Way Anova were used for continuous variables depending on data normality.

3. SRC-1 acts as a transcriptional co-repressor of the luminal marker, CD24 in endocrine resistance

3.1 Introduction

3.1.1. Global analysis of SRC-1 activity in endocrine resistant breast cancer

The basis for this research is founded on important preliminary work conducted by a previous postgraduate student Dr Damian McCartan. Given the interest of this research group in the SRC-1 protein, specifically in its associations with endocrine resistant breast cancer and the relative scarcity of human tumour tissue from resistant patients, large scale discovery experiments were pursued in the *in vitro* model of endocrine resistance.

A SRC-1 ChIP sequencing experiment was conducted to identify the global impact of SRC-1 at the genomic level in response to Tamoxifen treatment in an established resistant environment. Chromatin immunoprecipitation was carried out on LY2 cells which were either untreated or Tamoxifen treated for 45 minutes. This time period was previously determined to be optimal for SRC-1 binding. The captured DNA was purified and isolated for sequencing by collaborators at the Conway Institute, University College Dublin. The SRC-1 ChIPseq data identified the regions of the genome where SRC-1 sat down upon and which of these regions were gene coding regions. Comparative analysis of the untreated versus the Tamoxifen treated LY2 cell DNA could identify specifically which gene coding regions were being regulated in response to Tamoxifen. Given the strong association between SRC-1 and the agonistic effects of Tamoxifen in breast cancer, the Tamoxifen treated data subset was of most interest.

The ChIPseq data did not disclose whether the SRC-1 associated genomic regions were in fact transcriptionally regulated by SRC-1. To clarify this, an accompanying Affymetrix Human Genome U133 Plus 2.0 DNA microarray was carried out using the same LY2 cell line. The Affymetrix U133 array analyses expression of over 47,000 transcripts using 25-mer oligonucleotides. The LY2s were either transfected with siRNA against SRC-1 or with a scrambled negative control. The transfection was done in duplicate and then cells were either control treated or Tamoxifen treated in the same manner as the ChIPseq experiment. Again, keeping the data sets relevant to Tamoxifen, the results of the treated microarray were used for comparative analysis with the ChIPseq data.

The resultant genes which were regulated were cross referenced with the ChIPseq genomic coding regions of interest. A definitive list of SRC-1 regulated genes was acquired and separated into genes which were potentially upregulated by SRC-1 and genes which were potentially downregulated by SRC-1. This data provides a wealth of information with regards to the activity of SRC-1 in endocrine resistant breast cancer. Data from this study published

in 2011 reported over 2,000 genes which were upregulated in response to SRC-1 expression and investigated ADAM22 as a specific mediator of disease progression in the endocrine resistant setting [172].

These preceding experiments have formed the foundation stone for this research. This work proposes to investigate the additional 1,061 genes which were directly downregulated by SRC-1 in the same model of breast cancer. This research has focused on these negatively regulated genes with a view to determining whether the renowned steroid receptor coactivator could actually have a serious role as a steroid receptor corepressor alongside its established role as a coactivator and furthermore to what end could SRC-1's ability to corepress specific genes impact on its tumourigenic agenda in breast cancer.

3.1.2. CD24: A potential SRC-1 target gene

CD24 was first identified in humans on the surface of B lymphocytes. Specifically, CD24 is expressed by pre-B lymphocytes and becomes lost during the maturation process to plasma cells. CD24 was attached to the cell membrane via a glycosylphosphatidylinositol (GPI) anchor (Figure 3.1) [207]. GPI anchors contribute to the overall organization of membrane bound proteins and are important in apical protein positioning. GPI-anchored proteins also play a critical role in a variety of receptor-mediated signal transduction pathways. Biochemical studies have demonstrated that glycosylation of CD24 is highly variable and cell type dependent [207].

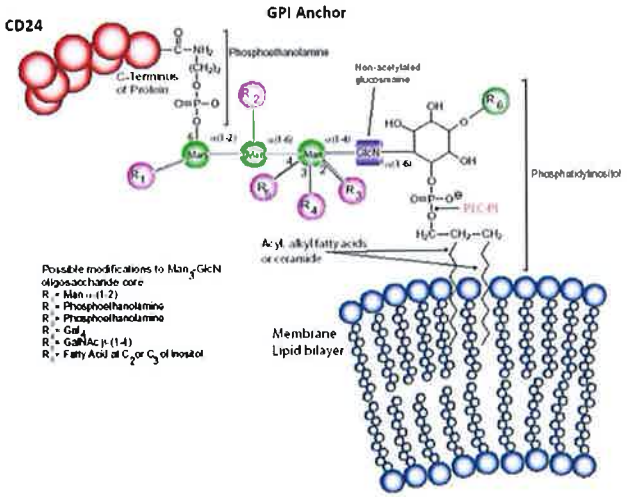


Figure 3.1. GPI structure of the CD24 protein. Glycosylphosphatidylinisotol (GPI) anchored proteins are linked at their carboxy terminus through a phosphodiester linkage of phosphoethanolamine to a trimannosyl-non-acetylated glucosamine (Man3-GlcN) core. The reducing end of GlcN is linked to phosphatidylinositol (PI). PI is then anchored through another phosphodiester linkage to the cell membrane through its hydrophobic region. *Adapted from www.sigmaaldrich.com.*

The human CD24 gene displays an allelic polymorphism (Val to Ala exchange) and is localised on the chromosome 6q21 [208]. CD24 knockout mice were found to be viable and displayed only a small block in B lymphocyte-poiesis suggesting that CD24 expression influenced the maturation of B-lymphocytes. Furthermore, overexpression of CD24 in lymphocytes suggested that CD24 levels can determine the capacity of early T and B lymphoid progenitors to proliferate and survive [209].

CD24 has been shown to have cell adhesion function and has been identified as a ligand for the $Ca^{(2+)}$-dependent lectin; P-selectin, which is expressed at the surface of activated platelets and endothelial cells. Binding of CD24 to P-selectin mediates leukocyte adhesion to vascular endothelium and platelets. CD24 was shown to support the adhesion of monocytes or neutrophils to activated endothelial cells or platelets both of which express P-selectin. These reactions could then be blocked by antibodies against CD24 or P-selectin [210]. It is unlikely that the CD24–P-selectin interaction accounts for all the biological activities that are attributed to CD24, since the protein is heavily dependent on glycosylation and is therefore subject to post-translational modification.

CD24 protein is also localised in lipid rafts at the cell membrane via its GPI-anchor [207]. Within rafts, CD24 is closely associated with src-kinases, G-proteins and calcium channels. Signal transduction via tyrosine phosphorylation and src-kinases has been described upon triggering of many GPI-anchored molecules therefore the absence or presence of CD24 may influence membrane raft composition and thereby affect important signalling pathways within the cell [211]. An association between CD24 and src kinases has been shown for small cell lung cancer and cross-linking of CD24 with specific antibodies resulted in tyrosine phosphorylation of cytoplasmic proteins in B lymphoblastic leukaemia [212]. Raft association may also play a role in the observed capacity of CD24 to regulate integrin activity.

3.1.3. CD24 expression in cancer

CD24's role in adhesion has aligned its expression with metastatic tumour progression. CD24 positivity has also been reported for a variety of the most common human cancers; haematological malignancies, ovarian cancer, renal cancer, prostate cancer, hepatocellular carcinoma, lung cancer, pancreatic cancer and breast cancer. Tumour metastasis requires tumour cells to shed off the primary tumour mass and invade the vasculature, before adhering again to the endothelia of target organs. It is further postulated that invading tumour cells in conjunction with platelets and fibrin form tumour thrombi in the bloodstream prior to invading the designated metastatic site. It is therefore conceivable that

CD24-positive tumour cells could attach to platelets and activate them via its established interaction with P-selectin and thus form part of the invasive tumour thrombi [210, 213]. This mechanism of action is further supported by animal models, in which either P-selectin deficiency or decreased levels of platelets have both been found to reduce haematogenous tumour cell metastasis [214].

3.1.4. CD24 expression in breast cancer

The role of CD24 in breast cancer has been quite contradictory. An immunohistochemistry study on human breast tumour tissue (n = 29) demonstrated strong CD24 expression in invasive breast cancer. An additional study reported membranous and cytoplasmic immunoreactivity in 85% of invasive breast cancer tissues. Furthermore, CD24 positivity was significantly associated with a positive nodal status, and showed a trend towards higher tumour grade. The CD24-positive tumours also associated with shorter disease-free survival times [215].

In contrast, some important studies have shown that loss of CD24 associates with tumour progression. CD24 expression was found to inversely correlate with the *in vitro* invasiveness of established breast cancer cell lines [216]. CD24 has become commonly associated with the cancer stem cell hypothesis ever since the pioneering work of AL-Hajj et al, in 2003 [217]. They identified a subset of breast cancer tumour stem cells which expressed high levels of CD44 and low levels of CD24. In the context of breast cancer, CD24 is now widely recognised as a marker of differentiated luminal cell types whilst the relative lack of CD24 in conjunction with CD44 is now strongly associated with a highly tumourigenic tumour cell phenotype [218]. Significantly, CD24 was identified as one of the genes associated with the luminal A signature in breast cancer subtyping from Sorlie, 2004 [16]. Another recent study from 2010 identified collectively lower levels of CD24 in the luminal B experimental cell lines and found that the CD24 levels correlated positively with the epithelial marker; E-cadherin and negatively with the basal marker, Vimentin. The study also showed that CD24 correlated tightly with many of the gene products that best distinguish luminal B cell lines from luminal A cell lines [219]. CD24 is also directly downregulated by estrogen and also by the oncogenic transcription factor, TWIST [220, 221]. As previously mentioned, TWIST is also a known target gene of SRC-1 [64]. Interestingly preliminary studies from Kaipparettu suggest that downregulation of CD24 could be regulated by epigenetic silencing including promoter methylation or histone deacetylase activity [220].

3.2. Aims

- To identify SRC-1 as a genuine corepressor in an endocrine resistant model of breast cancer.

- To investigate the specific gene, CD24 as a target of this direct and negative regulation by SRC-1 in an endocrine resistant model of breast cancer.

- To identify potential transcription factor hosts for SRC-1 in this context.

3.3. Work leading to the research hypothesis

3.3.1. The LY2 breast cancer cell line are resistant to the anti-proliferative effects of Tamoxifen

The LY2 cell line has been used to characterise and investigate the role of SRC-1 in endocrine resistant breast cancer. These cells were originally derived from the parental MCF-7 cell line and initially were developed as a resistant cell line to the antiestrogen, LY 117018. As a consequence of this, the LY2 cells developed a robust secondary resistance to the anti-proliferative effects of Tamoxifen treatment.

The MCF-7 and LY2 cell lines are used extensively in this research as a comparative model of endocrine sensitivity versus endocrine resistance. A proliferation assay was conducted to confirm the responsiveness of MCF-7s and LY2s to various endocrine treatments (Figure 3.2). Whist the colourimetric MTS assay is a common choice for determining drug cytotoxicity, the assay proved too insensitive for this course of treatment and so the rate of proliferation was calculated by manual cell counts using trypan blue dye and a haemocytometer. Both cell lines were steroid depleted for 72 hours to minimize any endogenous endocrine activity, the treatments were then applied for a further 72 hours before cells were trypsinised and counted. The endocrine sensitive cells exhibit increased rate of growth in response to estrogen treatment, this response is significantly curtailed in the LY2 cells. The MCF-7 cells also exhibit reduced rate of growth when treated with the anti-estrogenic Tamoxifen. In contrast, Tamoxifen increases proliferation in the LY2 cells thus confirming an innate resistance to the effects of Tamoxifen treatment and confirming the appropriate inclusion of the LY2 cell line as a model of endocrine resistant breast cancer.

As previously discussed in section 1.1.3, breast cancer subtypes are associated with significant prognostic value. Each cancer is characterised in terms of receptor expression (ER, PR, HER2), tumour grade and in some cases are also classified according to a Ki67 proliferation index. According to ATCC, both the MCF-7 and LY2 cells are ER positive and HER2 negative however they differ in regards to PR expression in that MCF-7 cells express it and LY2 cells do not. Both cells lines are therefore typically luminal type cancers. This fact was further confirmed using the hallmark luminal marker, E-cadherin and the basal counterpart, Vimentin. Western blot analysis for protein expression of these two markers was carried out on the MCF-7 and LY2 cells (Figure 3.3A, Figure 3.3B). Both cell lines were positive for E-cadherin and negative for Vimentin. In addition, the two cell lines were also analysed for EpCAM (Epithelial cell adhesion marker) cell surface expression using FACS

analysis (Figure 3.3C). Using mean fluorescence intensity, a quantitative value was assigned to EpCAM expression in each cell line and reported reduced levels of EpCAM in the LY2 cells. Although the LY2 cells are representative of a luminal cancer subtype, stratifications can be drawn between them and the MCF-7 cells. The LY2 cells exhibit reduced expression of important luminal genes, PR and E-cadherin and are an overall more aggressive cell line than the MCF-7s thus the LY2 cells are more likely aligned with a luminal B cancer subtype whilst the MCF-7 cells appear typically luminal A.

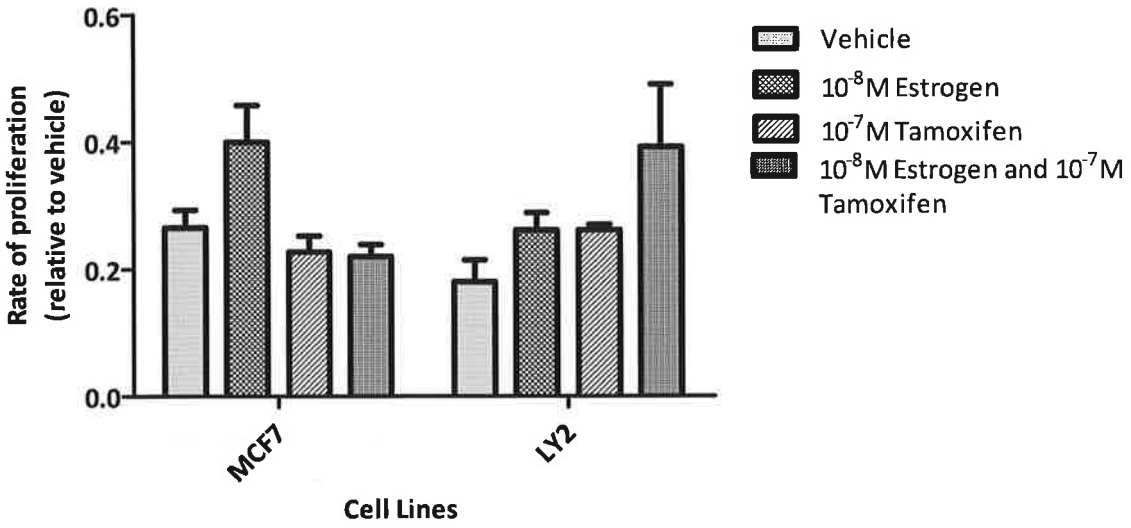


Figure 3.2. Characterisation of the MCF-7 and LY2 cell lines as an *in vitro* model of endocrine resistance. (A) A proliferation assay to determine sensitivity to Tamoxifen treatment of both cell lines. Cellular proliferation was assessed by manual cell counts following a 72 hour period of various endocrine treatments. The LY2 cells exhibit reduced responsiveness to the proliferative effects of estrogen when compared to the endocrine responsive MCF-7 cells. The LY2 cells also exhibited resistance to the anti-proliferative effects of Tamoxifen and instead showed a certain amount of proliferation in the presence of estrogen and Tamoxifen when compared to the MCF-7 cells. Results are expressed as mean ± SEM, n=4.

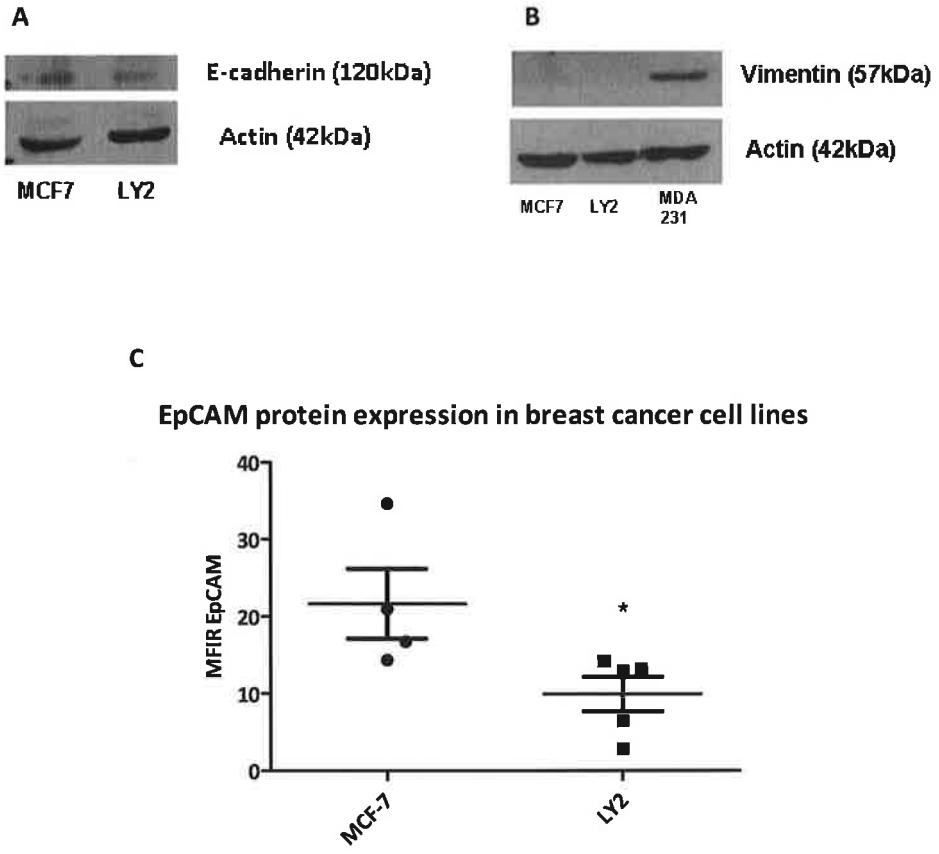


Figure 3.3. MCF-7 and LY2 cells represent luminal breast cancer subtypes

(A) To confirm the luminal nature of the breast cancer cell lines, MCF-7 and LY2 cells were lysed and run on an 8% polyacrylamide gel for the luminal marker, E-cadherin. The gels were transferred and probed with a mouse monoclonal antibody against E-cadherin. Western blot analysis showed E-cadherin was expressed in both cell lines however the LY2 cells exhibited slightly less of the protein compared to the MCF-7 cells. The membrane was subsequently reprobed for β-actin to confirm proper loading of each protein in the experiment. **(B)** The same cell lines were further analysed for expression of the basal cell marker Vimentin. Cell lysate from the MDA-231 cells were included in the experiment as a positive control for Vimentin. The cell lysates were run on a 10% polyacrylamide gel and probed with a mouse monoclonal antibody against Vimentin. Both MCF-7 and LY2 expressed no Vimentin thus confirming their luminal status. **(C)** FACS analysis of a second epithelial cell marker, EpCAM in MCF-7 and LY2 cell lines. EpCAM protein expression also associates with luminal breast cancer subtypes. The endocrine sensitive MCF-7 cells exhibit increased amounts of EpCAM expression compared to the LY2 cells.

3.3.2. SRC-1 expression associates with an endocrine resistant breast cancer phenotype

In addition to being resistant to the therapeutic effects of Tamoxifen, the LY2 cell line has elevated levels of SRC-1 expression compared to the MCF-7 cell line, and therefore makes these two cell lines a relevant and appropriate system in which to study the functionality of SRC-1 in endocrine resistant breast cancer.

Under basal conditions, RNA was extracted from MCF-7 and LY2 cell lines and SRC-1 gene expression was analysed via quantitative real time PCR (Figure 3.4A). The data obtained confirmed that the unstimulated LY2 cells expressed significantly higher levels of SRC-1 mRNA than the unstimulated MCF-7 cells ($p<0.05$).

One of the serious clinical dilemmas associated with Tamoxifen resistance is that not only do cells become resistant to the therapeutic benefits but that Tamoxifen itself can actually serve to advance the progress of the tumour within the breast. As previously described, the p160 family of proteins are potential targets of aberrant Tamoxifen action and it has been demonstrated that Tamoxifen can upregulate SRC-1 in a resistant breast cancer phenotype [167]. To confirm that this is also the case in our experimental *in vitro* model, SRC-1 was transiently overexpressed in MCF-7 cells (Figure 3.4B). The SRC-1 overexpressing cells were then treated with Tamoxifen and a significant upregulation of SRC-1 mRNA expression was observed after 4 hours ($p<0.05$).

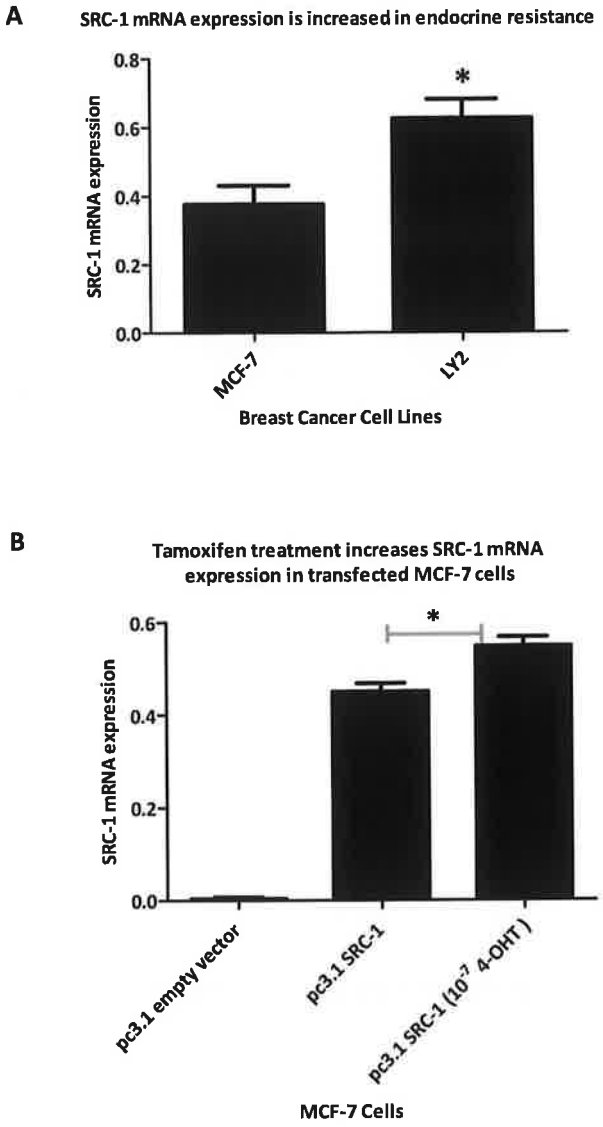


Figure 3.4. SRC-1 mRNA expression is upregulated in endocrine resistance and in response to Tamoxifen. (A) Quantitative real time PCR analyses of relative SRC-1 mRNA levels in the endocrine sensitive MCF-7 cells and the endocrine resistant LY2 cells. Results are expressed as mean ± SEM, n=3 (*p<0.05). **(B)** Quantitative real time PCR analyses of relative SRC-1 mRNA levels in MCF-7 cells which have been transiently transfected with the pc3.1 empty plasmid vector or the SRC-1 overexpression plasmid, pc3.1-SRC-1 or the pc3.1-SRC-1 plasmid treated with 10^{-7}M 4-OHT for 4 hours. Tamoxifen treatment significantly increases SRC-1 mRNA transcription. Results are expressed as mean ± SEM, n=3 (*p<0.05).

3.3.3. SRC-1 associates specifically with poor prognosis in luminal B breast cancer patient subpopulation

Previous work conducted by this research group has associated SRC-1 with an endocrine resistant phenotype using both the *in vitro* model system and a constructed tissue microarray from our clinical patient database. Using Kaplan Meier survival analysis, SRC-1 was identified as an independent predictor of disease free survival in the Tamoxifen treated patient population [169],

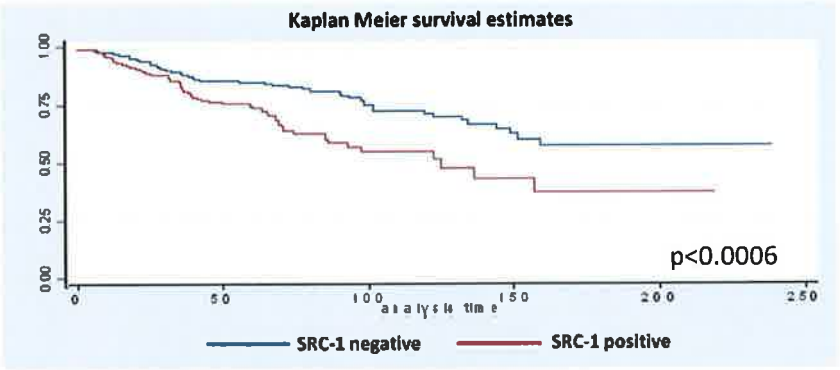
Using the same clinical patient population, stratifications were drawn in order to reclassify the patient population in terms of the four main breast cancer subtypes; luminal A, luminal B, HER-2 overexpressing and basal. This classification system is the result of global molecular profiling and distinct gene signatures have been assigned to each individual subtype (Table 3.1).

SRC-1 expression was assessed in terms of the breast cancer luminal subtypes in the clinical patient population (Figure 3.5). Kaplan Meier survival analysis significantly associated SRC-1 with poor disease free survival in the luminal A ($p=0.0006$) (Figure 3.5A) and the luminal B ($p=0.0009$) (Figure 3.5B) breast cancer subtypes in the patient population. The survival analysis identified a slight decrease in disease free survival in the luminal B patients who were positive for SRC-1 compared to the luminal A patients however additional statistical analysis of the TMA data using an odds ratio further confirmed the poor prognostic relationship between SRC-1 and the luminal B patients (Figure 3.5C). It has been suggested that the luminal B subtype reflects a group of patients who will not benefit from adjuvant Tamoxifen despite positive receptor values [222] therefore this clinical observation strengthens the evidence for SRC-1's specific function in endocrine resistant breast cancer.

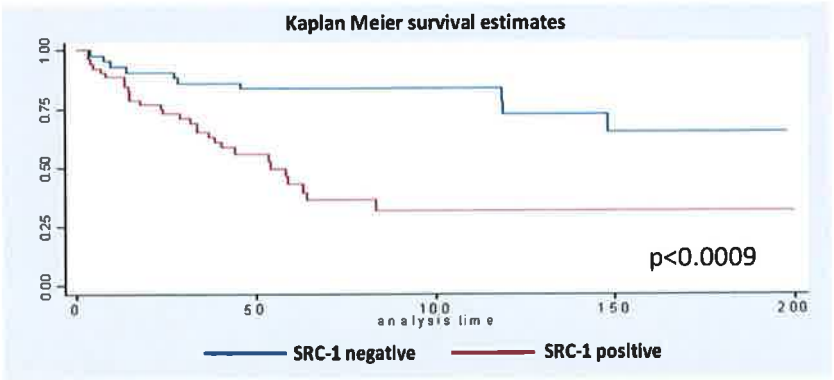
Luminal A	Luminal B	HER2 overexpressing	Basal-like
High ER/PR positivity	Low to moderate ER/PR expression HER expression	ER/PR negative HER2 positive	ER/PR/HER2 negative
High luminal gene expression	Low to moderate luminal gene expression High expression of novel gene cluster	High expression of genes in the HER2 amplicon	Negative for luminal gene expression High expression of basal gene cluster
Good disease free survival	Reduced disease free survival	Poor disease free survival	Poor disease free survival
	High expression of proliferative genes		High expression of proliferative genes

Table 3.1. General characterisation of the four major molecular subtypes of breast cancer. *Adapted from Sorlie, 2004* [16].

A Luminal A patients



B Luminal B patients



C

	Odds Ratio	p-value	Odds ratio 95%CI
SRC-1	1.81	0.001	1.4 – 2.4
SRC-1 & Luminal A	1.24	0.180	0.9 – 1.7
SRC-1 & Luminal B	2.60	0.001	1.4 – 3.9

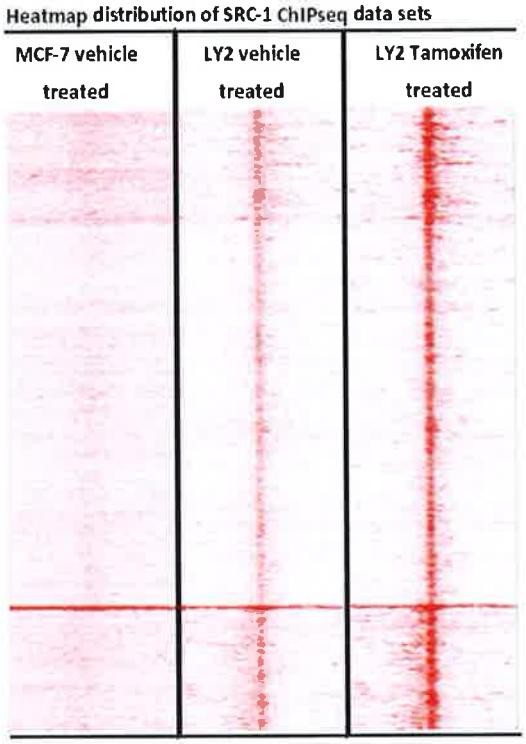
Figure 3.5. Increased SRC-1 expression associates specifically with a luminal B breast cancer subtype. Kaplan Meier estimates of disease-free survival according to SRC-1 expression in two common breast cancer patient subtypes; **(A)** luminal A patients and **(B)** luminal B patients (n=403, p<0.0006 and n=98, p<0.0009 respectively). **(C)** The odds ratio associated with SRC-1 expression in the luminal A and luminal B subtypes. The odds ratio suggests that the risk of poor disease free survival is significantly higher for patients who are positive for SRC-1 expression and who have luminal B breast cancer than if they had SRC-1 and luminal A breast cancer.

3.3.4. Global analysis confirms SRC-1 activity is specific to an endocrine resistant breast cancer phenotype

Despite the accumulation of evidence which has associated SRC-1 activity with the development of endocrine resistant breast cancer, little is known as to which signalling pathways or genes are being targeted by SRC-1 in this specific tumourigenic context. To address these questions, global analysis of SRC-1 activity was conducted by our research group in order to identify the specific regions of the genome where SRC-1 was having its most prominent effect. A chromatin immunoprecipitation sequencing (ChIPseq) experiment was conducted using the LY2 cell line. Subsequent sequencing data obtained from the SRC-1 ChIPseq experiment was merged with published data from another SRC-1 ChIPseq experiment that was conducted in MCF-7 cells (Figure 3.6A) [199]. Both data sets were visualised on a heat map to demonstrate that SRC-1 binding is much more prominent in the LY2 cell line compared to the MCF-7cell line. Furthermore, the distribution of SRC-1 peaks in the Tamoxifen treated LY2 sample is significantly more than the vehicle treated LY2 sample which indicates that SRC-1 activity at the genome is directly driven by treatment with Tamoxifen.

Additionally, the increased SRC-1 activity in the presence of Tamoxifen was also clearly represented when the number of SRC-1 ChIPseq peaks were counted (Figure 3.6B). This Tamoxifen associated increase was maintained even when a false discovery rate of <1 was applied to the experimental data. Importantly, the majority of peaks obtained from the SRC-1 ChIPseq experiment in the LY2 cells also occurred at functionally relevant genomic regions such as upstream of the transcription start site or within the first exon (Figure 3.6C).

A



B

SRC-1 peaks identified	Vehicle treated LY2s	Tamoxifen treated LY2s
Total	33544	51617
FDR<=1	299	20486

C

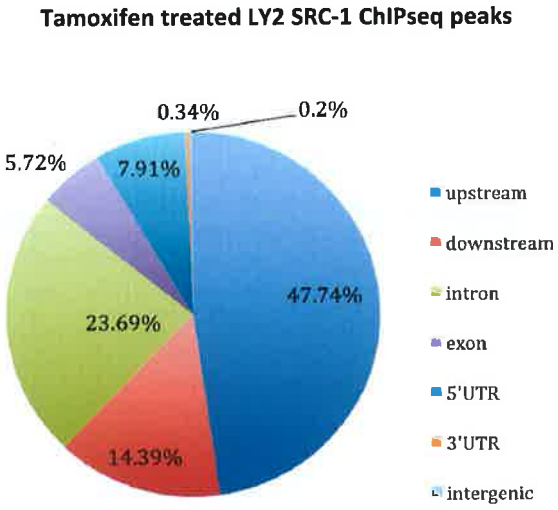


Figure 3.6. SRC-1 exhibits heightened genomic activity in endocrine resistance and in response to Tamoxifen. (A) A heatmap displaying clustering of SRC-1 ChIPseq datasets for vehicle treated MCF-7 cells, vehicle treated LY2 cells and Tamoxifen treated LY2 cells. Clusters are generated according to SRC-1 peak enrichment at the promoter regions. **(B)** Breakdown of the identified peaks from ChIP-sequencing of SRC-1 immunoprecipitated DNA, with a false discovery rate (FDR) of <1. **(C)** Pie-chart illustrating the genomic location of SRC-1 interaction sites following Tamoxifen treatment.

3.4. Results

3.4.1. SRC-1 directly downregulates a cohort of genes in endocrine resistant breast cancer

As previously discussed, the combined data sets from the SRC-1 ChIPseq and the SRC-1 DNA microarray produced two specific gene lists; genes which appeared to be upregulated and genes which appeared to be downregulated. Of particular interest to this research were the 1,061 genes which were significantly downregulated by SRC-1 (Figure 3.7). This gene list was composed from a comparative analysis of SRC-1 peaks which were identified either 5,000 base pairs upstream of the transcriptional start site or within the first exon using a false discovery rate of <1 (10,136 peaks) and the 2.047 genes which were identified from the DNA microarray as being upregulated in the Tamoxifen treated LY2 cells in the absence of SRC-1. The 1,061 genes represent the genes which were common to both data sets and it is this list of genes which is the basis of the hypothesis predicting SRC-1's corepressor activity. In order to assess what the predominant functionality of these 1,061 suppressed genes, pathway analysis was conducted on the identified genes using GeneGo software. The 10 most significant pathways were reported in the form of a bar chart. Apoptosis and cellular differentiation were highlighted as these pathways are of particular interest with regards to breast cancer and SRC-1 activity.

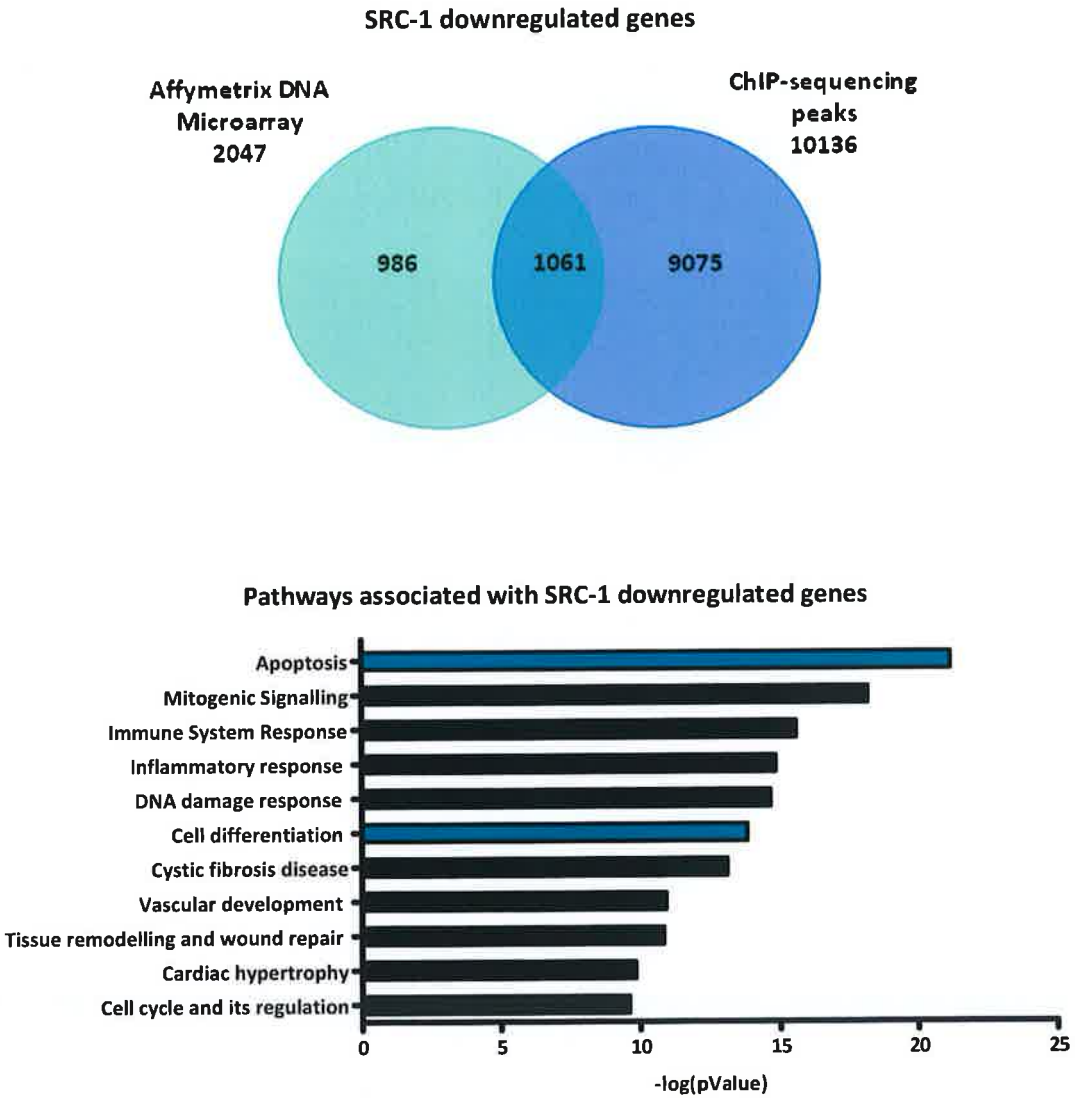


Figure 3.7. SRC-1 directly corepresses a cohort of target genes. The Venn diagram depicts the 2,047 genes which were upregulated in the absence of SRC-1 from the DNA expression microarray data set and the 10,136 SRC-1 peaks which were identified on the human genome by the SRC-1 ChIP-sequencing experiment. There are 1,061 overlapping genes which are downregulated by SRC-1 and which also have SRC-1 binding within 5,000 base pairs upstream of the gene promoter or within the first exon of the gene. A bar graph to represent pathway analysis associated with the 1,061 SRC-1 downregulated genes. Data is represented in terms of statistical significance.

3.4.2. The luminal marker CD24 is identified as a potential target gene for SRC-1 mediated corepression

Additional bioinformatic analysis was used to look at which specific genes were most significantly downregulated by SRC-1 in the LY2 cell line. The 1,061 genes were subsequently ranked in terms of fold enrichment and statistical significance (Figure 3.8). The 25 most significantly regulated genes are represented in a radial chart. CD24 was ranked number 21 out of the list of 1,061 genes. With regards to breast cancer, CD24 is well established in its role as a marker of luminal differentiation and it is also commonly used to characterise and define cancer stem cell populations. Given that luminal B cancers are often associated with reduced expression of luminal gene signature and a less differentiated cancer phenotype, it was of interest to investigate the ability of SRC-1 to suppress this well known luminal marker. Also pathway analysis in Figure 3.7 highlights cell differentiation as a significant pathway associated with SRC-1 action in the resistant model of breast cancer. For these reasons, CD24 was pursued as a breast cancer related gene of interest and was the subject of subsequent molecular investigation with regards to its relationship with SRC-1.

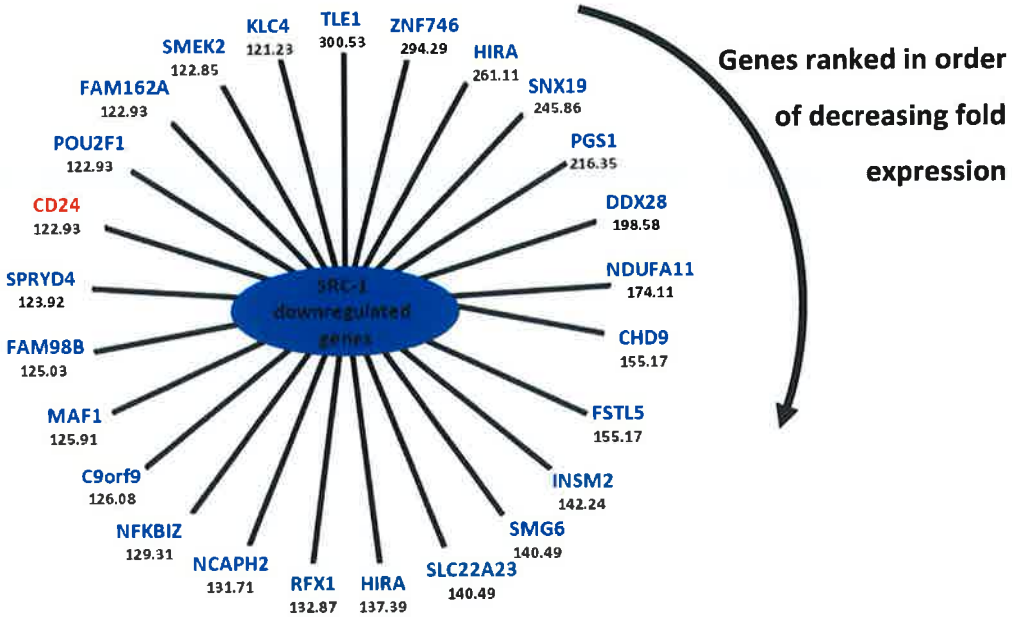


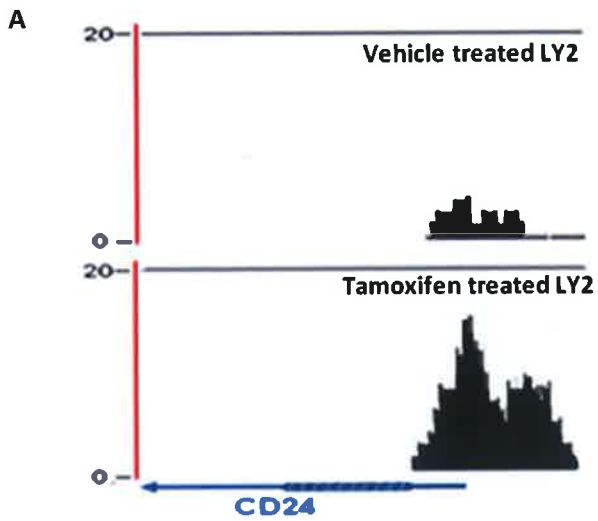
Figure 3.8. The luminal marker CD24 is identified as a potential target gene for SRC-1 mediated corepression. A radial diagram represents the top 25 genes which are directly downregulated by SRC-1 in the LY2 model of endocrine resistance. The genes have been ranked in terms of fold enrichment and the luminal marker CD24 (ranked number 21) is highlighted as a breast cancer related gene of interest on this list. *The genes are listed by gene name and associated fold enrichment number.*

3.4.3. SRC-1 is recruited to the CD24 promoter region in endocrine resistant breast cancer

As shown, the global analysis of SRC-1 activity has focused this research specifically on the genes whose expression appeared to be lost or suppressed as a result of SRC-1 signalling. From this data set, the CD24 gene has been selected to further investigate the hypothesis that SRC-1 has the capacity to dictate bidirectional signalling of different cohorts of target genes. CD24 expression and regulation has therefore been subjected to further molecular investigations to confirm the results inferred by the bioinformatic data. By demonstrating that SRC-1 can negatively regulate the CD24 gene will serve as a proof of principle for the proposed ability of SRC-1 to function both as a coactivator and as a corepressor to mediate tumour progression in endocrine resistance.

Analysis of the ChIPseq data reported recruitment of SRC-1 to the CD24 promoter. Images from the UCSC genome browser visualise the SRC-1 peaks present at the CD24 promoter region (Figure 3.9A). There is an obviously larger peak present in the Tamoxifen treated LY2 sample than in the comparative vehicle treated LY2 sample.

To further confirm the ChIPseq analysis, the ChIP experiment was repeated in the LY2 cells using the same treatment (Figure 3.9B) conditions as the original ChIPseq discovery experiment. Primers were designed against the CD24 promoter region and quantitative real time PCR was used to assess recruitment of SRC-1 to the CD24 promoter. The results showed that SRC-1 was recruited to the CD24 promoter over and above the recruitment observed at the IgG level.



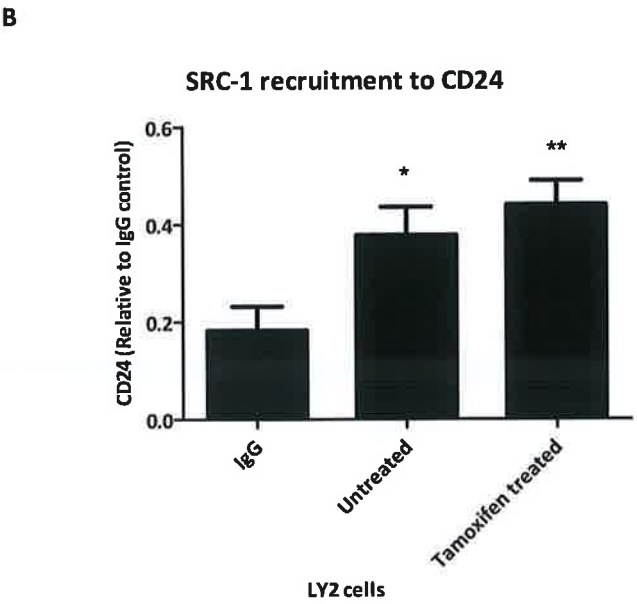


Figure 3.9. SRC-1 regulation of CD24 in the endocrine resistant LY2 cells. (A) Recruitment of SRC-1 to the CD24 promoter region is depicted by a representative image from UCSC genome browser. Comparisons are made between peaks under vehicle treated conditions and following treatment with Tamoxifen. **(B)** ChIP analysis of SRC-1 recruitment to the CD24 promoter was conducted to confirm the SRC-1 ChIPseq data. Quantitative real time analysis was used to quantify SRC-1 recruitment to the DNA and data is represented relative to the IgG negative control (n=4). Each sample is normalised to the specific experimental replicates.

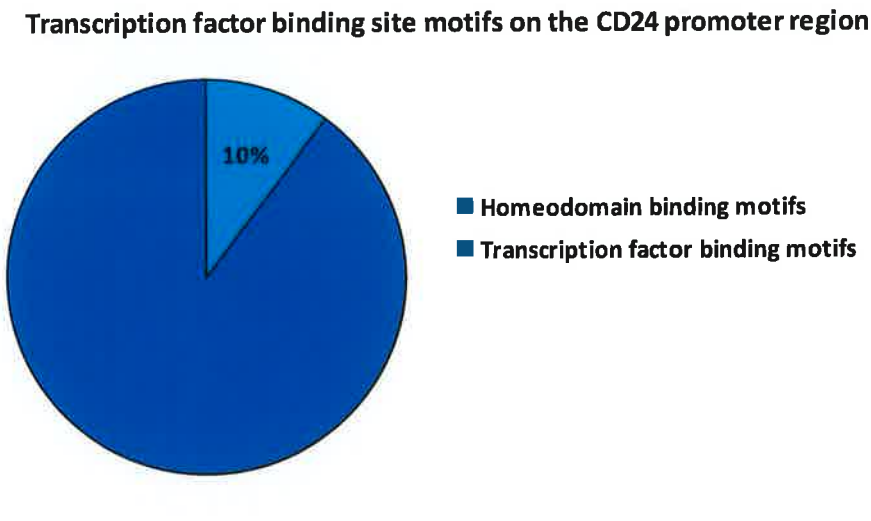
3.4.4. HOXC11 is a potential transcription factor for CD24 regulation by SRC-1

Although SRC-1 is a master regulator of gene expression, it is unable to bind directly to the DNA. Instead SRC-1 recruits in various transcription factors and additional transcriptional machinery to regulate target gene expression at the targeted DNA site. In the case of CD24, SRC-1 must coregulate its expression with a specific transcription factor. In order to identify potential transcription factors which may be working with SRC-1 to facilitate the transcriptional silencing of CD24, MatInspector software was used to search the CD24 promoter region for transcription factor binding site motifs.

An area of 3,000 base pairs upstream of the CD24 transcription start site was selected and subsequently scanned for binding site motifs. Within this region, 779 motifs were identified and of this 779, 79 (10%) were members of the various homeobox families (Figure 3.10). The homeobox superfamily of genes encode for a wide range of transcription factors which are commonly defined as master regulators of development due to their ability to activate or repress downstream target genes [223].

Previous work from this group identified the homeobox protein, HOXC11 as a functional partner for SRC-1 in endocrine resistant breast cancer [62]. HOXC11 was significantly upregulated in the LY2 cells (Figure 3.11A). Strong associations were also observed between HOXC11 and SRC-1 in the endocrine resistant LY2 cells compared to the MCF-7 cells. Importantly, similar observations were made in the equivalent clinical patient populations[62]. Furthermore, coexpression of HOXC11 with SRC-1 was noted as a superior predictor of poor clinical prognosis than any of the other classic parameters [62]. This previous relationship between SRC-1 and HOXC11 in conjunction with the abundance of homeobox binding motifs present on the CD24 promoter made HOXC11 a likely transcription factor candidate for SRC-1-mediated repression of CD24 expression. Subsequent sequence analysis of the CD24 promoter region confirmed the presence of two potential HOXC11 binding site motifs on the reverse strand (Figure 3.11B).

A



B

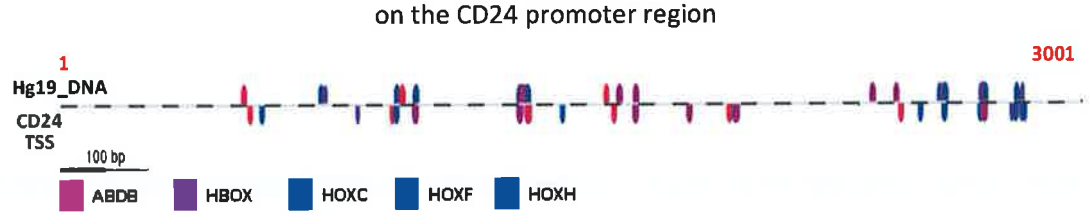
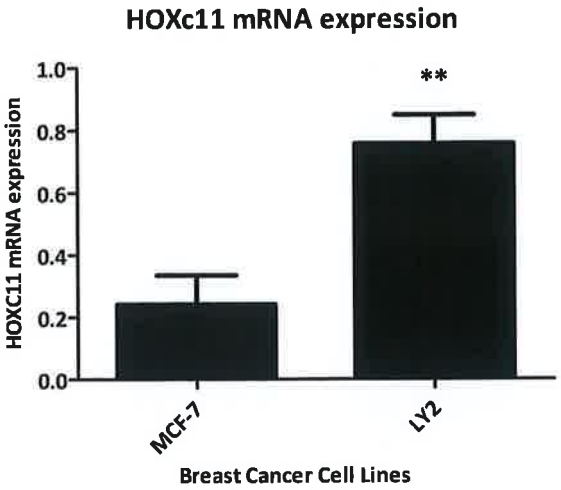


Figure 3.10. HOXC11 is a potential transcription factor partner for SRC-1 mediated downregulation of CD24. (A) MatInspector analysis of the CD24 promoter region identified 779 transcription factor binding site motifs within a 3,000bp region upstream of the CD24 transcription start site. There were 79 (10%) of the transcription factor binding site motifs which were matched to various homeobox family members. **(B)** A schematic image from the MatInspector software is included to visualise the homeobox family matrices associated with the CD24 promoter; ABDB, HBOX, HOXC, HOXF and HOXH.

A



B

HOXC11 binding site motif on the CD24 promoter region

Two found on the reverse strand:

Chromosome	Start	End	Strand	Score	Sequence
Chr. 6	21,155,076	21,155,087	-	0.75	GCAGCCATAAAA
Chr. 6	21,155,527	21,155,539	-	0.72	TGAGTCATTAAC

Figure 3.11. HOXC11 has two specific binding motifs within 1000bp of the CD24 transcriptional start site. (A) Quantitative real time PCR analysis shows increased HOXC11 mRNA expression in the LY2 endocrine resistant cells compared to the MCF-7 cell line,(n=4, $p<0.01$). **(B)** The CD24 promoter was scanned for HOXC11 transcription factor binding site motifs and two such sites were located on the reverse strand of the CD24 promoter.

3.4.5. HOXC11 is recruited to the CD24 promoter with Tamoxifen treatment

Although investigation of the CD24 promoter region did show two potential transcription factor binding motifs for HOXC11, it was important to determine whether or not HOXC11 was capable of regulating CD24 transcriptional activity in the same context as SRC-1.

To this end, a RNA sequencing experiment which had been previously conducted by Dr. Damian McCartan was reanalysed specifically for CD24. RNA sequencing is used to analyse gene expression under specific experimental conditions. In this instance, LY2 cells were transfected with either scrambled control or with siRNA against HOXC11 and the resultant mRNA content of each sample was sequenced to determine which genes HOXC11 was affecting in endocrine resistance. Total RNA was isolated from the cells and enriched for mRNA by targeting the 3' polyadenylated tail. The mRNA was then fragmented and reverse transcribed to synthesise cDNA. The cDNA was then sequenced in a high throughput fashion to generate a HOXC11 specific transcriptome. Access to this data, was useful in determining whether or not HOXC11 appeared to impact on CD24 transcriptional activity. Images of the LY2 samples with and without HOXC11 expression were depicted using the UCSC genome browser (Figure 3.12A). The peaks correlate to transcriptional activity at the CD24 gene site. There are peaks present in both LY2 samples however there appears to be an increase in the intensity of the peaks in the absence of HOXC11 which supports the hypothesis for HOXC11s' participation in the repression of CD24 transcription.

To further confirm the regulation of CD24 by HOXC11, the original ChIP experiment was repeated using a HOXC11 antibody and the same conditions as were reported for the SRC-1 ChIPseq (Figure 3.12B). The results showed that HOXC11 was recruited to the CD24 promoter over and above the recruitment observed at the IgG level in the vehicle treated sample and further again in the Tamoxifen treated sample. This data in conjunction with the SRC-1 ChIP from Figure 3.9 suggest that both SRC-1 and HOXC11 are recruited to the CD24 promoter in an *in vitro* model of endocrine resistance.

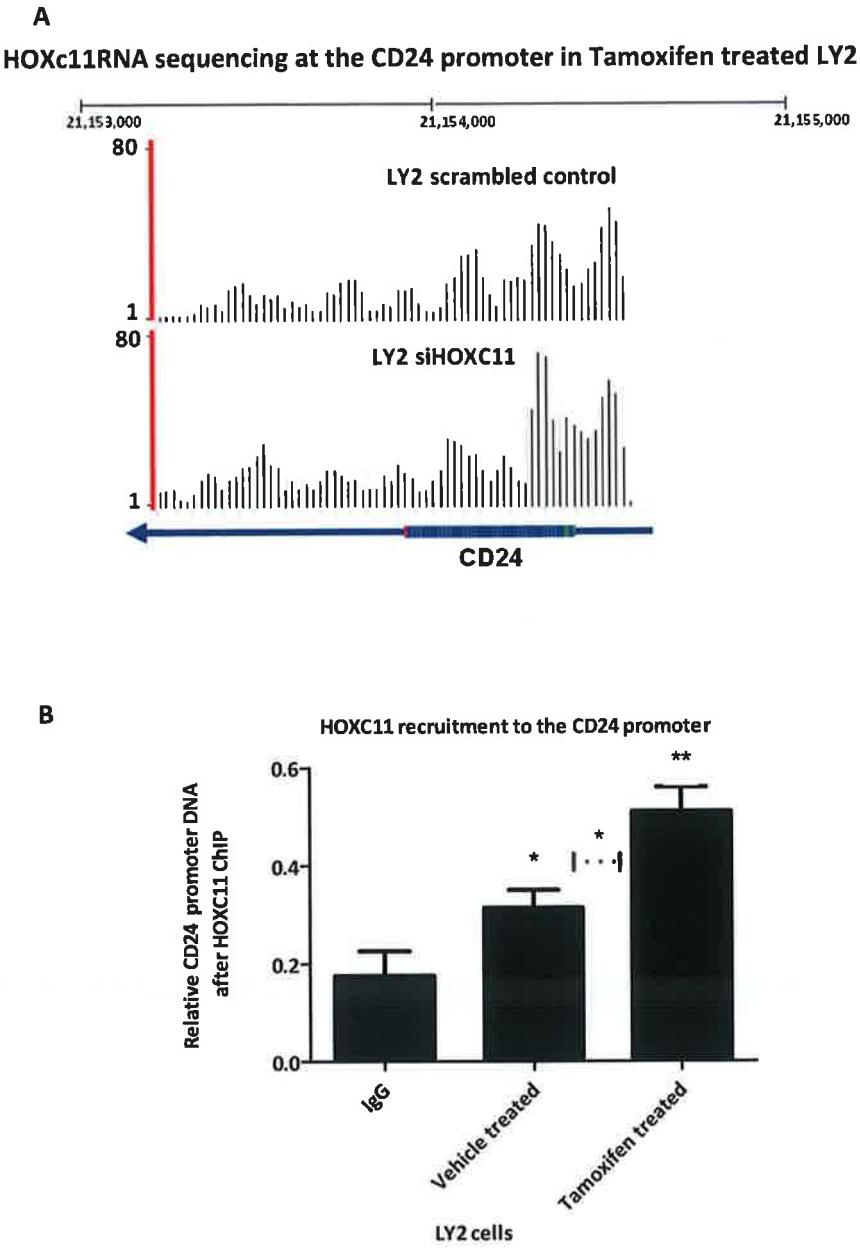


Figure 3.12. HOXC11 regulation of CD24 in the endocrine resistant LY2 cells.

(A) HOXC11 RNAseq in LY2 cells shows transcriptional regulation of CD24 cDNA as depicted by a representative image from UCSC genome browser. **(B)** ChIP analysis of HOXC11 recruitment to the CD24 promoter was conducted to confirm the recruitment of HOXC11 to the same promoter region as SRC-1. Quantitative real time analysis was used to quantify HOXC11 recruitment to the DNA and data is represented relative to the IgG negative control (n=3). (*, $p<0.05$ and **, $p<0.01$ relative to the IgG control). Each sample is normalised to the specific experimental replicates.

3.4.6. CD24 expression is reduced in the endocrine resistant model of breast cancer

CD24 has been proposed as a direct downstream target of SRC-1 negative regulation as a result of a series of discovery experiments. However, it was important to assess the expression profile of CD24 in the *in vitro* experimental model using untreated endocrine sensitive MCF-7 cells and untreated endocrine resistant LY2 cells.

Given the strong associations between SRC-1 and the endocrine resistant phenotype, it was expected that CD24 expression would be lost as a tumour cell progresses from an endocrine sensitive state towards an endocrine resistant state. Quantitative real time PCR analysis of both the MCF-7 cells and the LY2 cells confirmed that there was significantly more CD24 mRNA expression in the MCF-7 cells than in the LY2 cells (Figure 3.13A). Furthermore, flow cytometry analysis of CD24 cell surface expression in the same cell lines confirmed that this observation held true at the protein level as well (Figure 3.13B). A representative histogram outlines the fluorescent output of CD24 positive stained cells over the negative PE-isotype control. The MCF-7 cells exhibit more CD24 staining than the LY2 cells. The FACS data was reanalysed in order to account for the intensity of CD24 staining in both cell lines using a mean fluorescence intensity ratio (MFIR) (Figure 3.13C). The MFIR is reported as a box plot and again confirms that the MCF-7 cells express more CD24 than the LY2 cells in a basal untreated state.

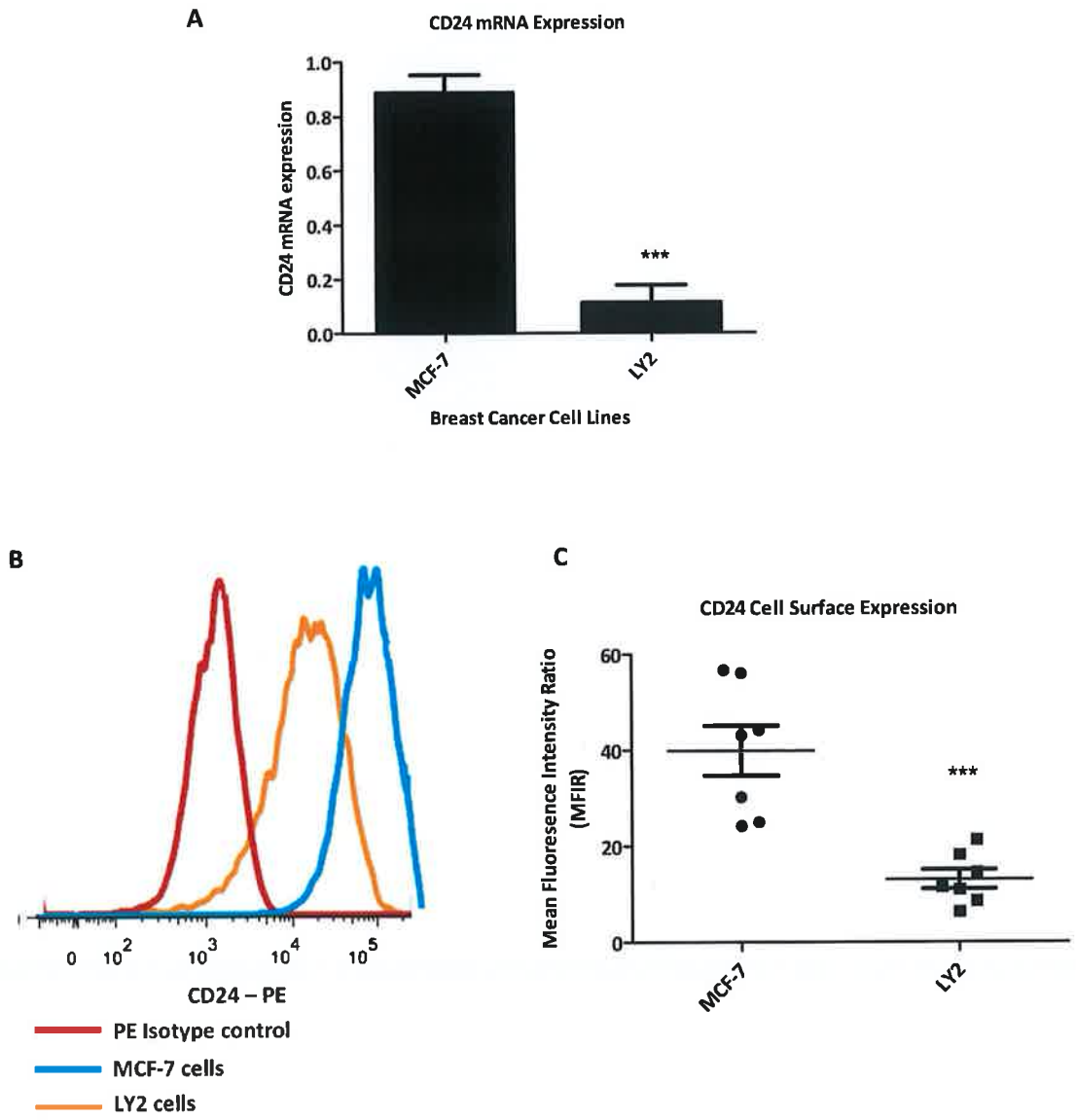


Figure 3.13. CD24 expression is lost in endocrine resistance. (A) Quantitative real time PCR analyses of relative CD24 mRNA levels in the endocrine sensitive MCF-7 cells and the endocrine resistance LY2 cells. Results are expressed as mean ± SEM, n=4 (***p<0.0001). **(B)** A representative histogram of CD24 cell surface expression in MCF-7 cells and LY2 cells as analysed by FACS flow cytometry. Cell surface expression is calculated as the percentage population minus PE isotype control. **(C)** A scatter plot of CD24 mean fluorescence intensity ratio (MFIR) in MCF-7 cells versus LY2 cells. The MFI of CD24 is calculated relative to the MFI of the specific PE isotype control for each cell line. (n=7, *** p<0.0001).

3.4.7. SRC-1 and HOXC11 negatively regulate CD24 expression in breast cancer cell lines

In order to confirm the regulatory role of SRC-1 and HOXC11 with regards to the expression of CD24, transient transfection experiments were carried out in the MCF-7 and LY2 cells (Figure 3.14). SRC-1 and HOXC11 gene expression was knocked out in the LY2 cells using predesigned siRNA gene silencing molecules. The LY2 cells were transiently transfected with the specific siRNAs for 72 hours after which approximately 80-85% of target genes expression was knocked down in both. Quantitative real time PCR analysis was used to determine any changes in CD24 mRNA expression in response to the transient knockdown of either SRC-1 or HOXC11. As expected, loss of SRC-1 in the LY2 cells induced an increase in CD24 mRNA expression (Figure 3.14A). A similar pattern was observed in the LY2s which had been transfected with siRNA against HOXC11 (Figure 3.14C); they also had increased levels of CD24 mRNA expression relative to the scrambled control.

To test the robustness of SRC-1 and HOXC11 regulation of CD24, an opposing experiment was carried out whereby the MCF-7 cells were transiently transfected with separate overexpression plasmids for both SRC-1 and HOXC11. Both overexpression experiments were conducted for 24 hours and quantitative real time PCR was used again to analyse any resultant changes in the transcriptional regulation of CD24. When either SRC-1 (Figure 3.14B) or HOXC11 (Figure 3.14D) were overexpressed in the endocrine sensitive MCF-7 cells, a concomitant decrease was observed in the gene expression levels of CD24.

In addition to the ability of SRC-1 and HOXC11 to regulate the transcriptional activity of CD24, it was important to investigate whether this regulation translated to the functioning of the CD24 protein in the endocrine resistant breast cancer model. Flow cytometry was used once more to study the impact on CD24 protein expression (Figure 3.15). The transient knockdown of SRC-1 was repeated again in the LY2 cells and the cells were stained for CD24 cell surface expression and analysed on the FACS Aria II flow cytometer. Loss of SRC-1 increased CD24 cell surface expression as depicted on the representative dot plot (Figure 3.15A) and histogram graphs (Figure 3.15B). The impact of SRC-1 on CD24 cell surface expression was also analysed using MFIR, the results of which concurred with the percentage of CD24 positive cells per cell population (Figure 3.15C).

The flow cytometry experiment was repeated for the HOXC11 knockdown in LY2 cells (Figure 3.16) however neither the percentage of cells stained positive for CD24 (Figure 3.16A) nor the intensity of the staining as measured by the MRIF (Figure 3.16B) identified any significant regulation of CD24 at the functional protein level by HOXC11.

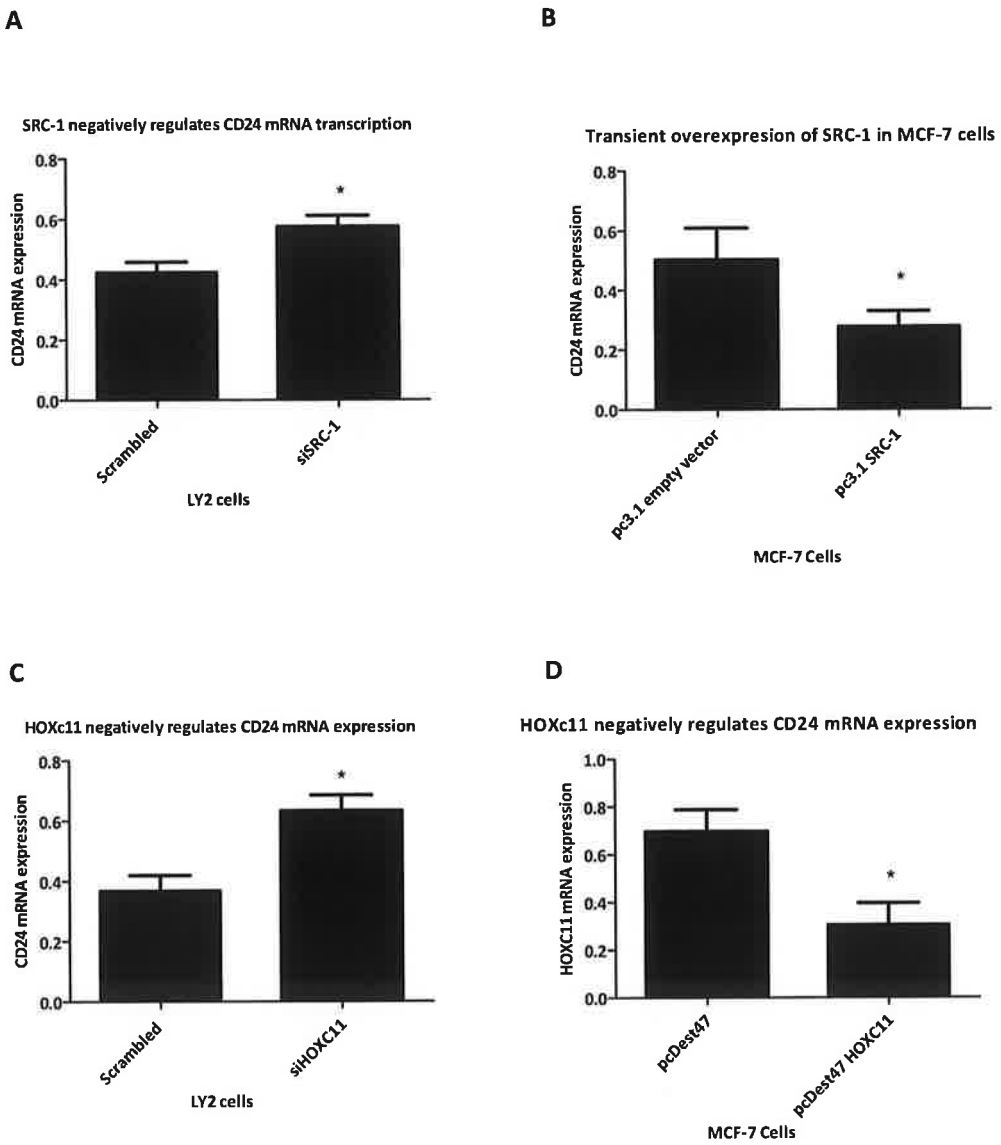


Figure 3.14. SRC-1 and HOXC11 negatively regulate CD24 transcription *in vitro*. Quantitative real time PCR analysis of CD24 mRNA expression in transiently transfected MCF-7 and LY2 cells. **(A)** SRC-1 negatively regulates CD24 mRNA expression when it is knocked out in LY2 cells (n=3, $p<0.05$) and **(B)** when it is overexpressed in MCF-7 cells (n=3). **(C)** HOXC11 also negatively regulates CD24 mRNA expression when it is knocked out in LY2 cells (n=3, $p<0.05$) and **(D)** when it is overexpressed in MCF-7 cells (n=3,*, $p<0.05$ relative to the scrambled control).

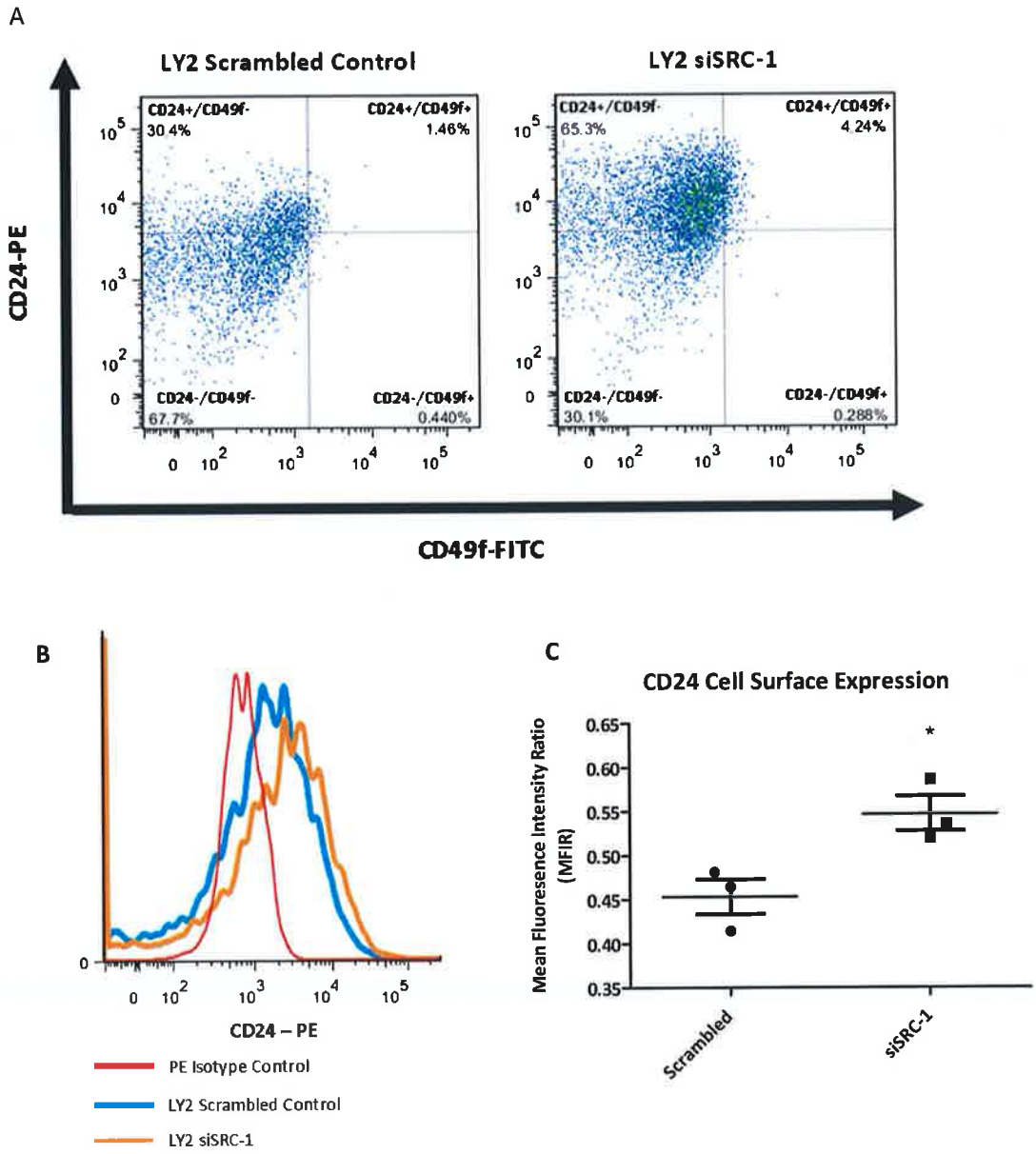


Figure 3.15. SRC-1 negatively regulates CD24 cell surface protein expression. (A) A representative flow cytometry profile of LY2 cells transiently transfected for 72 hours with scrambled control or siRNA against SRC-1. The cells are analysed for CD24 and CD49f cell surface expression. Loss of SRC-1 results in an upward shift in CD24 cell surface expression. **(B)** A corresponding histogram highlights the increase in CD24 cell surface expression in the absence of SRC-1 expression relative to the PE isotype control. **(C)** A scatter plot demonstrates an increase in CD24 MFIR in LY2 cells which were transiently transfected with siRNA against SRC-1 for 72 hours. The MFIR results are normalised and graphed relative to experimental replicates (n=3, *, $p<0.05$).

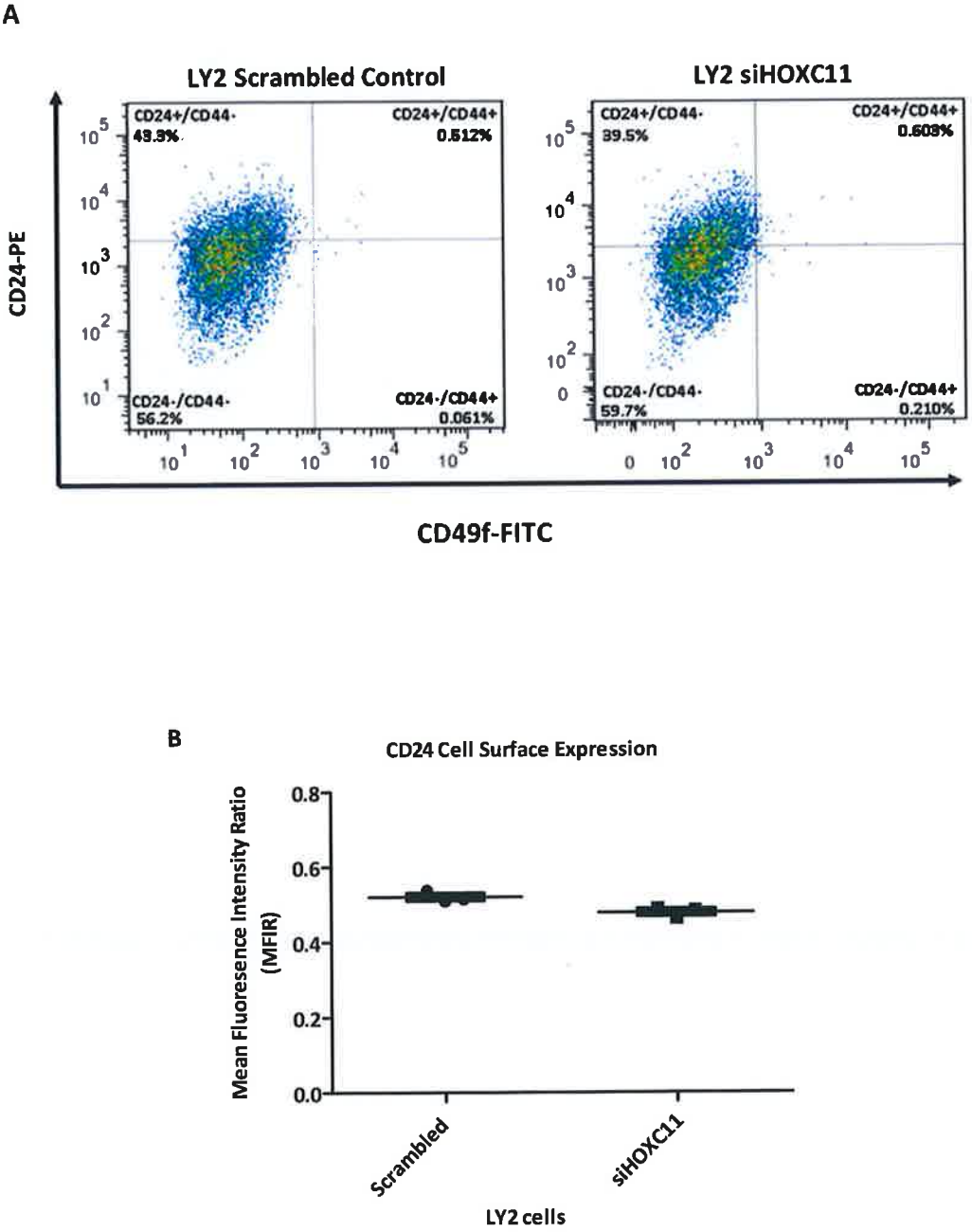


Figure 3.16. HOXC11 does not significantly regulate CD24 cell surface expression. (A) A representative flow cytometry profile of LY2 cells transiently transfected for 72 hours with scrambled control or siRNA against HOXC11. The cells are analysed for CD24 and CD49f cell surface expression but there is no obvious change in the CD24 profile when HOXC11 is knocked out. **(B)** A scatter plot shows the MFIR of CD24 in the same LY2 cells. Again, there is no discernible difference observed between the two experimental samples (n=3).

3.4.8. CD24 is negatively regulated by SRC-1 *in vivo*

Due to a long standing collaborative relationship with Prof. Jianming Xu at Baylor College of Medicine, Houston, it was possible to verify the effect of SRC-1 on CD24 using their *in vivo* murine model. The Xu lab has made significant contributions to the field of steroid receptor coactivator research, and in recent years the lab has engineered a SRC-1 knockout *MMTV-PyMT* (mouse mammary tumour virus-polyoma middle T) transgenic mouse. The MMTV-PyMT mice form spontaneous mammary gland tumours by 5 weeks of age and have a high incidence of lung metastasis. One of the major observations from the *SRC-1*$^{-/-}$ *MMTV-PyMT* model was that loss of SRC-1 significantly reduced breast cancer metastasis to the lung. Using RNA isolated from both wildtype *MMTV-PyMT* and *SRC-1*$^{-/-}$ *MMTV-PyMT* mammary gland tumours, CD24 expression was analysed using quantitative real time PCR analysis (Figure 3.17). As seen in the *in vitro* model system, CD24 mRNA expression was upregulated in the absence of SRC-1 activity.

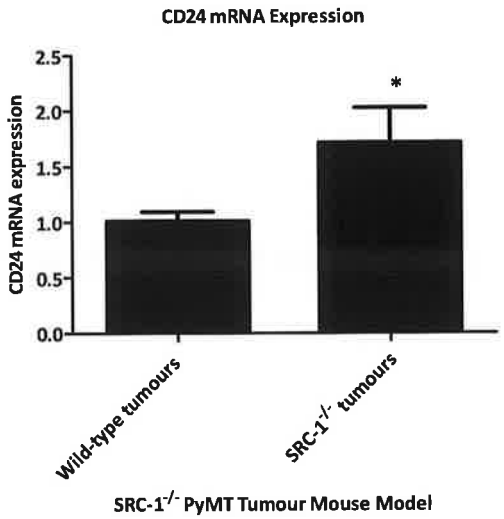


Figure 3.17. SRC-1 negatively regulates the transcription of CD24 mRNA *in vivo*.

Quantitative real time PCR analysis of CD24 mRNA expression in mammary gland tumours from *MMTV-PyMT* transgenic mice. CD24 mRNA expression is increased in the *SRC-1*$^{-/-}$*;MMTV-PyMT* mice in which the SRC-1 gene has been genetically silenced compared to the *wildtype MMTV-PyMT* mice (n=4, *, $p<0.05$).

3.5. Discussion

Endocrine resistance is a complex clinical phenomenon which impacts on a reported 40% of the breast cancer patients, undergoing prescribed targeted therapy. Given the relative dominance of Tamoxifen as a targeted endocrine therapy in breast and the longevity of its use in the clinic, resistance to this specific drug is a common occurrence. Unfortunately, as endocrine resistance is associated with an advanced and increasingly aggressive tumour course, access to endocrine resistant breast cancer tissue can prove quite difficult and therefore a robust *in vitro* model is all the more important for clinical research efforts. The LY2 cells in conjunction with the MCF-7 cells have been used by this research group for a number of years and have provided a stable and well established *in vitro* model of Tamoxifen resistance. The LY2 cells were originally derived from the MCF-7 cell line thus the differences between the two cell lines can largely be attributed to exogenous factors such as their differing culture conditions and treatment regimes. The LY2s are consistently maintained in low levels of 4-OHT and are resistant to the antagonistic effects of Tamoxifen treatment (Figure 3.2). Importantly, the LY2 cells are not simply impassively resistant to Tamoxifen, they have concomitantly developed a response to the agonistic effects of Tamoxifen treatment and therefore can actually proliferate and promote tumour progression in its presence. The LY2 cells although they still retain hormone receptor expression and luminal classification (Figure 3.3), exhibit a reduced responsiveness to steroid signalling (Figure 3.2). Additional previous experiments conducted by this group have shown that LY2 cells also express increased levels of growth factor signalling molecules such as Ets2, c-myc, MMP9 and src kinases when treated with Tamoxifen [112, 113]. This observation in conjunction with the observed resistant phenotype of the LY2 cells also makes them an applicable model of luminal B breast cancers in accordance with the classifiers outlined in Table 3.1.

The continued investigation of the steroid receptor coactivator SRC-1 and its activity in breast cancer is the corner stone of this research. Important work from the SRC-1 knockout mouse models confirmed that SRC-1 is not a dominant force in events which determine tumour initiation or growth but instead mediates tumour metastasis and therefore its expression is often a hallmark of tumour progression. As a breast cancer develops and gains momentum, regulated steroid signalling within the tumour mass is lost and receptor switching between the steroid and growth factor pathways becomes far more commonplace. As a result, treatment options become far more limited. The main biological function of SRC-1 is to mediate successful and appropriate transcription reactions of specific

NRs such as ERα, thus the observed deregulation of ERα signalling in tumour cells goes hand in hand with deregulation of SRC-1 activity in these same cells. For these reasons, aberrant SRC-1 activity has been strongly associated with the development and maintenance of endocrine resistant tumours, both *in vitro* and *in vivo* [74]. This observation of aberrant expression is further corroborated by whole genome analysis of SRC-1 which clearly shows that SRC-1 activity is increased in the resistant breast cancer cells, specifically in response to Tamoxifen treatment, compared to the endocrine sensitive breast cancer cells (Figure 3.6A). Research efforts into SRC-1 have uncovered its ability to bind a diverse range of transcription factors and coactivate multiple genes in multiple biological contexts. Moreover, it has highlighted the inappropriate hyperactivity of SRC-1 in cancer. A number of SRC-1 target genes have already been identified with regards to metastatic spread however little is known about the genes or pathways targeted by SRC-1 in an endocrine resistant state [74]. Whole genome analysis of SRC-1 when combined with the SRC-1 knockout DNA microarray produced copious amounts of data with regards to potential and novel target genes of SRC-1 activity. An interesting development was the significant number of genes which appeared to be negatively regulated by SRC-1; this data essentially redefines the accepted view of SRC-1 as a renowned coactivator and suggests that it may also function as a corepressor. Further analysis of this dataset showed that the luminal marker CD24 was in the top 2% of significantly regulated genes (Figure 3.8). Furthermore, the identification of a HOXC11 binding site on the CD24 promoter (Figure 3.10) immediately warranted a more detailed investigation into the regulation of CD24. Investigation of CD24 in the *in vitro* model of resistance confirmed its negative regulation by SRC-1 and HOXC11 at the transcriptional level (Figure 3.14) and by SRC-1 at the protein level (Figure 3.15). Unfortunately, a knockdown of HOXC11 alone was not sufficient to induce loss of CD24 at the protein level. It is possible that SRC-1 has other mechanisms of negatively regulating CD24 separate to its interactions with HOXC11 which would explain why a knockout of SRC-1 can induce an increase in CD24 whereas a knockout of HOXC11 cannot (Figure3.16). CD24 has two distinct ERE binding sites and has previously been shown to be negatively regulated by ER in breast cancer cells therefore it is possible that SRC-1 could also be coactivating ER-mediated suppression of CD24 activity [220]. In addition, another study associated Twist protein with negative regulation of CD24 in breast cancer [221]. Twist is a known target of SRC-1 activity and its expression is aberrantly upregulated in breast cancer [64]. These studies support the possibility that there is a functional redundancy with SRC-1 mediated suppression of CD24.

CD24 has been a marker of well differentiated tumour cells and is often a feature of luminal A type tumours. The molecular characterisation of the major breast cancer subtypes was based on 552 genes, 22 of these genes were overexpressed in the luminal A signature, of which CD24 was one [9]. In addition, a study by Blick et al, 2010, conducted genome wide transcriptional profiling on numerous breast cancer cell lines to investigate the relationship between EMT and breast cancer stem cells. Their data confirms the transcriptional suppression of CD24 in LY2 cells and also positively correlated CD24 with the epithelial marker E-cadherin and negatively correlate it with the basal cell marker, Vimentin.
The most prominent functional role of CD24 has been its coexpression with cell surface markers such as CD44 or CD49f in order to profile and characterise tumour cell populations. In this context loss of CD24 is associated with the highly tumourigenic and chemo-resistant cancer stem cell or basal cell population. These CD24 negative populations exhibit a high capacity for tumourigenesis and metastasis. In this context, CD24 is an obvious target choice for SRC-1 activity in endocrine resistance. An inappropriate rise in SRC-1 generally coincides with a progressively deregulated cellular state. Loss of luminal features is characteristic of this progression therefore SRC-1 targeting of luminal markers such as CD24 for corepression could be a key manoeuvre in the shift from a responsive to a resistant state.

4. Identification of PAWR, as an additional SRC-1 / HOXC11 target gene in endocrine resistant breast cancer

4.1. Introduction

4.1.1. Combined global analysis of SRC-1 and HOXC11 in endocrine resistant breast cancer

It has been reported that both coregulators and homeobox proteins have the ability as "master regulators" to regulate gene expression upwards or downwards depending on the environmental context [177, 224]. It is thus hypothesised that the complex formed between SRC-1 and HOXC11 can switch its functionality at differing gene sites in order to maximise its tumourigenic impact as a typically luminal tumour cell gains independence from steroidal signalling mechanisms. The relationship between SRC-1 and the non steroidal transcription factor HOXC11 has already been identified as a feature of endocrine resistance both *in vitro* and *in vivo* [62]. Moreover, the current molecular evidence suggests that this relationship can function to increase the transcriptional activation of S100β, and decrease that of the luminal marker, CD24 in the same model of endocrine resistance.

Given the interesting findings surrounding the SRC-1 and HOXC11 complex in endocrine resistance, further molecular investigations were undertaken to identify additional downstream targets. Given that the SRC-1 ChIPseq data was already available, the most comprehensive approach to identifying novel downstream targets of SRC-1 and HOXC11 was to conduct a HOXC11 ChIPseq experiment in the same cells and under the same experimental conditions. Just as the SRC-1 ChIPseq had been, the HOXC11 ChIP was also carried out by Dr. Damian McCartan and the sequencing was outsourced to collaborators at UCD. The two datasets were then analysed for common binding sites. HOXC11 had significantly less of a genomic imprint than that of SRC-1 in the LY2 cells, however the pattern of regulation remained the same as there was increased HOXC11 binding in the presence of Tamoxifen compared to the untreated LY2 sample. Using the previous FDR of <1 for the SRC-1 ChIPseq and a FDR of <10 for the HOXC11 ChIPseq, bioinformatic analysis identified a total of 54 common regions, 32 of which were known protein coding regions.

4.1.2. PAWR; a second target of SRC-1/HOXC11 mediated repression

Of the 32 potential target genes, the PAWR (PKC apoptosis WT1 regulator; also known as PAR-4, (prostate apoptosis response-4)) gene was ranked 2nd in terms of statistical significance. There was no accompanying DNA microarray data for the HOXC11 ChIPseq experiment therefore it was unclear as to which direction the potential 32 target genes were being regulated. Literature searches were conducted on the 32 genes and a small number of cancer related genes were selected for further investigation. PAWR was one of

these genes due to the fact that it has anti-apoptotic function and is a well characterised tumour suppressor gene in prostate cancer.

Human PAWR was discovered by yeast two hybrid studies as a partner of Wilms' tumor 1 (WT1) [225] and atypical protein kinase C (aPKC) [226]. The PAWR gene is located on the minus strand of chromosome 12q21.2 and is evolutionarily conserved in vertebrates [227]. The PAWR protein contains conserved amino acid residues on its SAC domain that are the target of phosphorylation by protein kinases A (PKA) and C (PKC), which regulate its ability to translocate to the nucleus and its ability to dimerize with other proteins [228]. The PAWR protein binds to and forms complexes with various proteins, including PKCÍ, WT-1, ZIPK, DAXX, and THAP1 through its C-terminal leucine zipper domain to affect cell survival [225, 229-231]. There is evidence that PAWR displays its pro-apoptotic activity by downregulating the anti-apoptotic Bcl-2 protein [232]. In addition, experimental studies showed that phosphorylated PAWR induces apoptosis through its ability to activate the pro-death, FasL-Faz-FADD-caspase 8 pathway and by inhibiting the NF-κB pro-survival pathway [230, 233].

In normal tissue, low expression of PAWR is found in certain terminally differentiated cells, such as neurons, lymphocytes, and epithelial cells of the mammary gland, ductal cells of the prostate, and smooth muscle cells. In contrast and consistent with its pro-apoptotic functions, high expression of PAWR levels are generally found in dying cells [234].

The sophisticated nature of PAWR anti-apoptotic activity is evident in neuron development, where apoptosis is fundamental for the control of the final number of neurons and glial cells in the central and peripheral nervous system. Excessive death of one or more populations of neurons results in disease or injury. PAWR is employed to mediate normal embryonic neuronal development, and prevents hyper-proliferation of nerve tissues. This is achieved by the asymmetric distribution of PAWR protein during the mitosis of neuronal progenitor cells; the daughter cells lacking PAWR differentiate into neurons, while those with high levels of PAWR undergo apoptosis [235].

4.1.3. Loss of PAWR activity in cancer

Tumour-suppressor genes play pivotal roles in maintaining genomic integrity and regulating cell proliferation, differentiation, and apoptosis, thus, loss-of-function mutations in these genes are directly related to tumourigenesis. Current molecular evidence has identified PAWR as a candidate tumour suppressor gene as downregulation of PAWR is seen in a variety of cancers, such as renal-cell carcinomas [236], neuroblastoma [237], acute lymphoblastic, leukemia, chronic lymphocytic leukemia [238] and endometrial cancer [239]. Evidence from PAWR knockout mice has shown that they develop spontaneous tumours in

various tissues, including endometrium, liver, and lung, and exhibit prostatic intraepithelial neoplasia (PIN). These mice also show an increased incidence of chemical- or hormone-inducible tumours of the bladder and endometrium [240]. Interestingly, heterozygous loss of PAWR yields the same frequency of tumours/PIN as homozygous loss. Overexpression studies with PAWR have shown to be sufficient to induce apoptosis in most cancer cells in the absence of a second apoptotic signal, but does not cause apoptosis in normal or immortalized cells [233, 241]. PAWR has also been found to be an essential downstream regulator of cell-death programs initiated by various exogenous signals, such as TRAIL, doxorubicin, and radiation [242].

There are various mechanisms which can induce a loss of PAWR expression and thus promote tumourigenesis. As previously mentioned, the PAWR gene is located in chromosome 12q21, this particular chromosomal region is quite unstable and is often deleted in pancreatic and gastric cancer[238, 243]. Although rare, PAWR expression can also be ablated due to a single base pair mutation. A recent study identified a single base mutation at aa189 in exon 3 of the PAWR gene, and an additional mutation at the effector domain of PAWR; which was associated with premature termination of PAWR in human endometrial carcinoma [239]. Interestingly, the same study also associated low levels of PAWR with PAWR promoter hypermethylation in 32% of analysed tumours. Silencing of PAWR via promoter hypermethylation was also detected in the endometrial cancer cell lines; SKUT1B and AN3CA [239]. Finally, AKT activity, which is commonly increased in cancer cells via growth factor signalling stimulation, oncogene activation, or loss of phosphatase and tensin homolog (PTEN) tumour suppressor activity, can also inhibit PAWR's pro-apoptotic function [244]. AKT binds to the PAWR protein and phosphorylates PAWR and sequesters it in the cytoplasm by the chaperone 14-3-3 proteins, thus inhibiting PAWR's apoptotic effects [245].

4.1.4. PAWR and breast cancer

There are very few studies that have looked at PAWR expression in breast cancer. Nagai et al, 2010, conducted the first large study of PAWR expression in primary breast cancer samples via immunohistochemical analysis of a tissue microarray dataset [246]. The results showed that loss of PAWR expression associated with breast cancer progression and also correlated closely with the increased expression of the growth factor receptors HER2 and EGFR as well as upregulation of phosphorylated AKT. The findings corroborate the observed functionality of PAWR in other cancer types and suggest that PAWR suppression or functional inactivation may protect breast tumour cells from cell death thus enhancing the

malignant phenotype and ultimately assisting a tumour cell in its attempts to escape targeted therapy.

Additional studies in MCF-7 breast cancer cells have shown that PAWR can be downregulated in response to estrogen and in response to insulin growth factor 1 (IGF-1) in this specific cell model [247]. IGF-1 acts primarily through binding to the IGF-1 receptor at the cell surface which can feed forward to activate the PI3-kinase signalling pathway, which regulates proliferation and, most importantly, resistance to apoptosis. Activation of the PI3-kinase pathway can then induce AKT to phosphorylate PAWR at the S249 residue, and render it inactive by sequestering it in the cytoplasm with 14-3-3 chaperone proteins. The downregulation of PAWR by estrogen appears to be mediated through ERα however the authors also suggest that the possibility of crosstalk between estrogen and IGF-1-mediated downregulation of PAWR, resulting in a synergistic increase in cellular proliferation and tumour progression [247].

4.2. Aims

- To identify new downstream target genes for the SRC-1/HOXC11 complex in endocrine resistance, by comparing the SRC-1 ChIPseq data with additional HOXC11 ChIPseq data.

- To confirm the binding of both SRC-1 and HOXC11 to the target gene *in vitro*.

- To determine the transcriptional consequence of SRC-1 and HOXC11 binding to PAWR *in vitro* and *in vivo*.

4.3. Results

4.3.1. Combined ChIPseq experiments identify common target genes of SRC-1/HOXC11 functional interactions in endocrine resistant breast cancer

As previously mentioned, HOXC11 was first identified from a mass spectrometry-based screen in the endocrine resistant LY2 cells [62]. Since this discovery, HOXC11 in conjunction with SRC-1 has proven to be a powerful predictor of disease free survival in Tamoxifen treated patients. Furthermore, S100β was identified as a target gene for SRC-1 and HOXC11 in this resistant context. Given the powerful clinical data supporting a genuine tumourigenic role for SRC-1/HOXC11 activity in endocrine resistance, it was important to pursue additional downstream targets and elucidate pivotal signalling mechanisms under the regulation of these two proteins together.

To identify such target genes in an unbiased manner, a HOXC11 ChIPseq experiment was conducted for subsequent comparison with the SRC-1 ChIPseq experiment (Figure 4.1). The HOXC11 ChIPseq experiment was conducted using the exact experimental conditions as the SRC-1 ChIPseq to ensure that direct comparisons could be made between the two. The sequencing was also carried out by the same collaborators therefore the resultant bioinformatic analysis of the two data sets could be applied to uncover common target genes which were regulated by both SRC-1 and HOXC11 in an endocrine resistant environment. The HOXC11 ChIPseq experiment produced significantly less peaks compared to that observed in the SRC-1 ChIPseq experiment and a different FDR was applied to its analysis (Figure 4.1A). Although, the number of peaks are low, the majority of them are situated either upstream or downstream of a gene start site which suggests that HOXC11 is binding to regulatory sites on the genome (Figure 4.1C). Furthermore, the average profile around the transcription start site and the transcription termination site shows clustering of HOXC11 binding peaks around the start site, indicating that these are likely to be sincere transcriptional targets (Figure 4.1D). The combined analysis of the two ChIPseq data sets produced 54 genomic peaks which were common to both proteins (Figure 4.1B).

A

	Vehicle treated	Tamoxifen treated
SRC-1 Intervals per sample	33,545	51,617
FDR < 1	289	20,486
HOXC11 Intervals per sample	663	1,558
FDR < 10	1	1,517

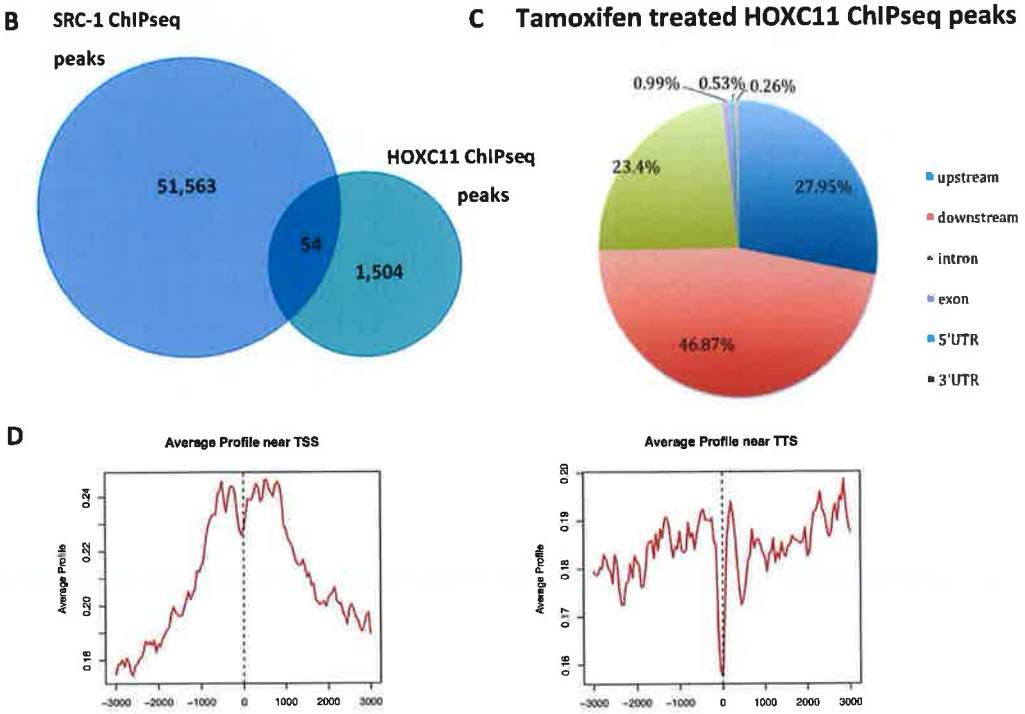


Figure 4.1. Combined global analysis of SRC-1 and HOXC11 activity in endocrine resistant breast cancer. The ChIPseq experiment was repeated in LY2 cells for HOXC11 using the same experimental conditions as the SRC-1 ChIPseq experiment ie 45 minute treatment with either vehicle or Tamoxifen. **(A)** Breakdown of the identified peaks from genome sequencing of SRC-1 immunoprecipitated DNA with a false discovery rate (FDR) of <1 and the identified peaks from the HOXC11 immunoprecipitated DNA with a false discovery rate (FDR) of <10. **(B)** Overlapping peaks between the SRC-1 ChIPseq data set and the HOXC11 ChIPseq data set. **(C)** Pie-chart illustrating the genomic location of HOXC11 interaction sites following Tamoxifen treatment; upstream and downstream simply refer to the gene start site. **(D)** Transcription start and termination site plot. The panels display the average ChIP enrichment signals surrounding the transcriptional start site (TSS) and the transcriptional termination site (TTS).

4.3.2. Identification of SRC-1/HOXC11 downstream target genes

Using bioinformatic tools, the 54 genomic peaks which were common to both SRC-1 and HOXC11 were further investigated to identify the regions beneath the peaks. Of the original 54 common peaks, 32 peaks represented gene coding regions and these peaks were subsequently ranked in terms of statistical significance (Figure 4.2). The anti apoptotic protein PAWR was ranked as the 2nd most significant among all of the 32 protein coding genes (Figure 4.2). The proteins are distributed among various chromosomal positions and literature searches identified a number of the 32 genes as having associations with cancer. Investigation of these novel target genes began with PAWR as it is a well known tumour suppressor gene in prostate cancer whereby it induces apoptosis in cancer cells but not in normal cells. In addition, apoptosis was cited as the most significant pathway in the pathway analysis of the 1,061 SRC-1 downregulated target genes in Figure 3.7, therefore PAWR was pursued with regards to a possible functional mechanism for SRC-1/HOXC11 activity in the endocrine resistant setting.

To confirm the transcriptional regulation of PAWR by HOXC11, the promoter region of the PAWR gene was scanned for transcription factor binding motifs for the HOXC11 protein (Figure 4.3). The same HOXC11 motif was used to scan the PAWR promoter as had been previously used to scan the CD24 promoter (Figure 3.11), the analysis confirmed that there were numerous HOXC11 binding sites present 1000bp upstream of the PAWR trasncriptional start site. Since SRC-1 is unable to bind directly to the DNA, it is likely that the SRC-1/HOXC11 complex utilises one of these non steroidal response elements to interact with the DNA to influence the transcriptional activity of PAWR. Specifically, there is one HOXC11 binding site located on the negative strand, which is a likely binding site as the PAWR gene is also located on the negative strand of chromosome 12. The confirmation and presence of HOXC11 binding site motifs further substantiates the accuracy of the ChIPseq experiment and the subsequent global analysis of HOXC11 activity in the endocrine resistant LY2 *in vitro* cell model.

Table 4.1. List of 32 protein coding genes identified from the combined analysis of the SRC-1 and HOXC11 ChIPseq experiments

	Chromosome	Gene Name	X.10.log10.pvalue.	Strand
1	chr1	CSDE1	969.71	-
2	chr1	GUCA2A	115.98	-
3	chr2	CALM2	481.27	-
4	chr2	ZAP70	144.61	+
5	chr3	CSRNP1	547.77	-
6	chr3	UBE2E2	347.2	+
7	chr3	PYDC2	218.76	+
8	chr5	TERT	164.29	-
9	chr6	RHAG	174.89	-
10	chr6	PBOV1	160.53	-
11	chr7	HSPB1	320.17	+
12	chr7	LRWD1	213.05	+
13	chr7	VOPP1	126.27	-
14	chr9	C9orf5	239.71	-
15	chr11	PHF21A	467.49	-
16	**chr12**	**PAWR**	**1257.43**	-
17	chr12	MAP3K12	207.72	-
18	chr12	HOXC6	116.69	+
19	chr13	GTF3A	157.23	+
20	chr14	TTC6	2189.94	+
21	chr14	MAX	401.36	-
22	chr14	NPAS3	242.96	+
23	chr16	GABARAPL2	243.44	+
24	chr17	CBX8	623.63	-
25	chr17	CEP95	584.79	+
26	chr17	HEATR6	356.32	-
27	chr17	MSI2	183.47	+
28	chr19	HNRNPM	183.22	+
29	chr20	NFATC2	228.51	-
30	chr21	TFF1	418.76	-
31	chr21	PSMG1	177.62	-
32	chr22	XBP1	567.96	-

Ranked in order of chromosomal number

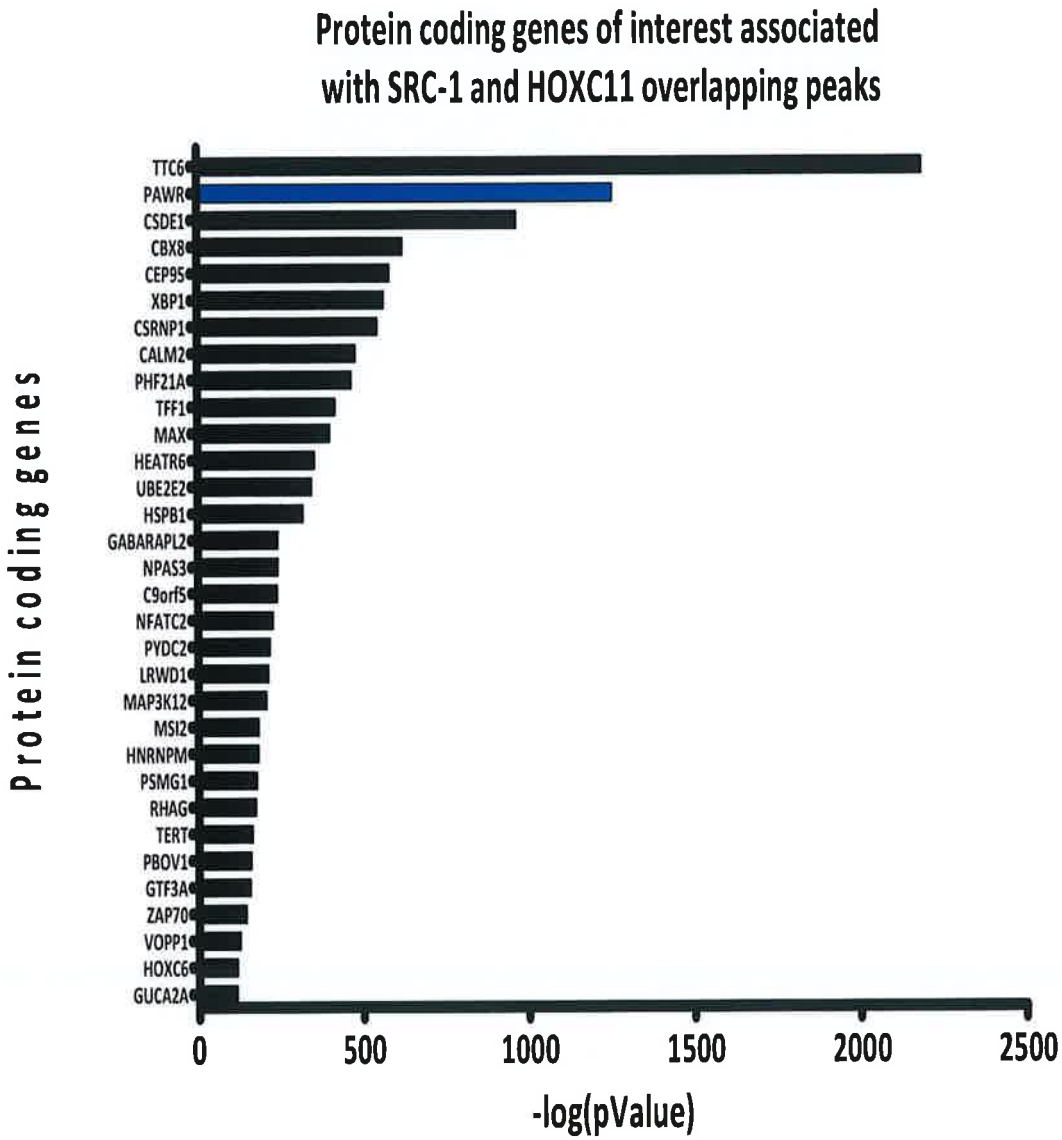
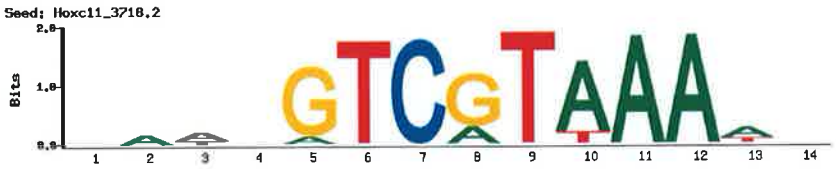


Figure 4.2. Bioinformatic analysis identifies the tumour suppressor gene PAWR as a significant dual target of SRC-1 and HOXC11 in endocrine resistant breast cancer. The SRC-1 ChIPseq (FDR <1%) overlapped with the HOXC11 ChIPseq (FDR <=10%) at 54 genomic regions, of which 32 were associated with protein coding genes. A bar chart shows these 32 genes ranked in order of statistical significance. The tumour suppressor candidate gene; PAWR ranked 2nd most statistically significant of all 32 protein-coding genes.

HOXC11 motif on PAWR upstream 1000 bp



HOXC11 binding site motifs found on the forward strand:

Chromosome	Start	End	Strand	Score	Sequence
Chr 12	148	155	+	0.65	ATCTCAAA
Chr 12	197	204	+	0.65	CTGCTAAA
Chr 12	435	442	+	0.65	ATCTGAAA
Chr 12	574	581	+	0.65	GTCTTGTA
Chr 12	579	586	+	0.66	GTACTTAA

HOXC11 binding site motifs found on the reverse strand:

Chromosome	Start	End	Strand	Score	Sequence
Chr 12	70	77	-	0.65	ATCTGAAA

Figure 4.3. HOXC11 has transcription factor binding sties on the PAWR promoter region.

HOXC11 has five specific binding motifs on the forward strand within 1000bp of the PAWR transcriptional start site and has one specific binding motif on the reverse strand of the same region.

4.3.3. PAWR is regulated by SRC-1 and HOXC11 in endocrine resistant breast cancer

Given the anti tumourigenic nature of PAWR in cancer, it was an interesting candidate gene for investigation for potential corepression by the SRC-1/HOXC11 complex. If SRC-1 and HOXC11 could downregulate genes in endocrine resistant breast cancer, then it is hypothesised that it is genes such as PAWR which would be ultimately targeted as the expression of such genes is in essence a barrier to tumour progression and metastasis.

Returning to both the SRC-1 ChIPseq raw data and the HOXC11 ChIPseq raw data, images of the representative peaks at the PAWR gene promoter site were captured using the UCSC genome browser (Figure 4.4). The differences observed in the recruitment of either protein to the PAWR promoter in the untreated LY2 sample compared to the Tamoxifen treated sample is generally indicative of regulation. Recruitment of SRC-1 to the PAWR gene promoter region appears to occur in both of the LY2 samples, however the peaks appear more intense in response to Tamoxifen treatment. There is significantly less HOXC11 activity at the PAWR promoter region when compared to that observed with the SRC-1 protein. However the pattern of regulation remains similar as again there is a much more substantial HOXC11 peak in the Tamoxifen treated LY2 sample compared to the untreated LY2 sample.

Although PAWR was initially identified as a result of bioinformatic analysis of the combined ChIPseq experiments, it was necessary to confirm any recruitment to PAWR by SRC-1 and HOXC11 by replicating the individual ChIP experiments in the laboratory and using quantitative real time PCR to detect any differences in recruitment between the samples (Figure 4.5). The PCR results from the SRC-1 ChIP experiment mimics the expected pattern of recruitment to the PAWR promoter in the LY2 cells however there is no significant difference with regards to the presence of Tamoxifen. The HOXC11 ChIP experiment shows significant recruitment of HOXC11 to the promoter region of PAWR over the IgG negative control in both the untreated and Tamoxifen treated samples. There was no difference in HOXC11 promoter occupancy that could be attributed to the presence of Tamoxifen.

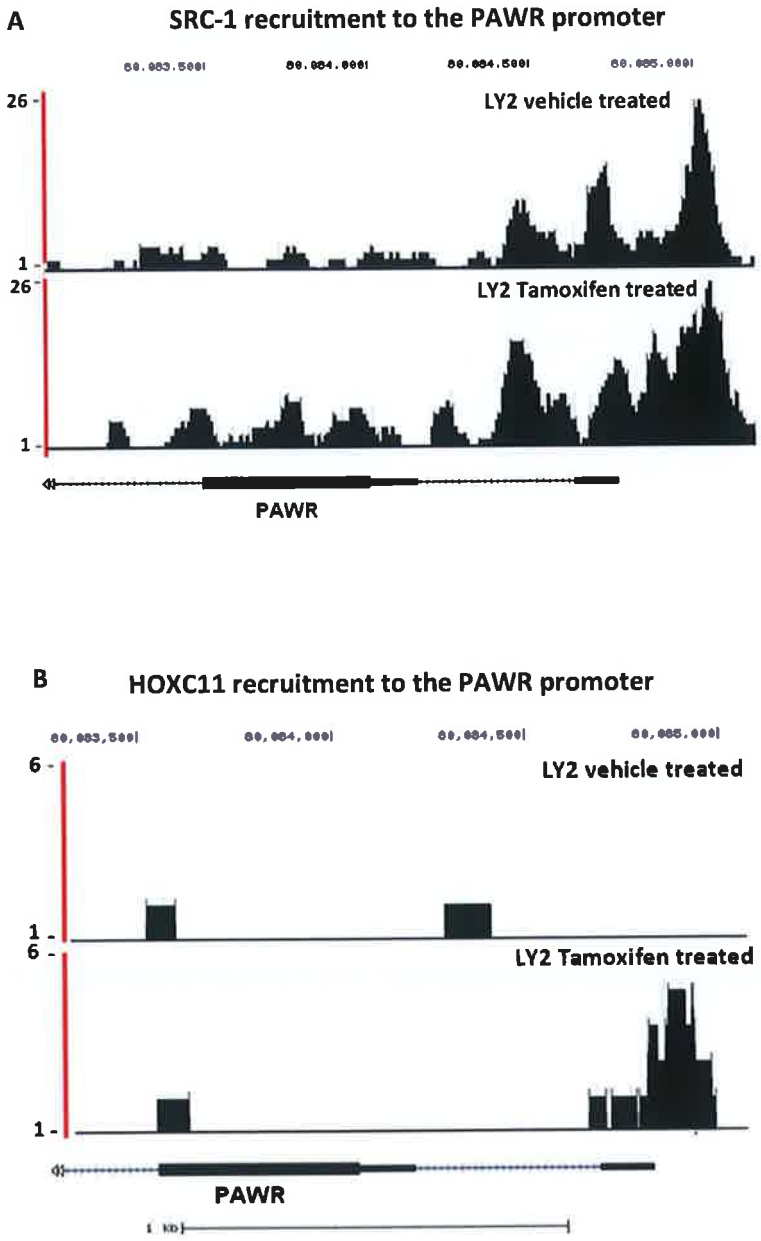


Figure 4.4. Occupancy of SRC-1 and HOXC11 at the PAWR promoter in LY2 cells. Representative images from UCSC genome browser of **(A)** SRC-1 ChIPseq and **(B)** HOXC11 ChIPseq data which show SRC-1 and HOXC11 occupancy at the PAWR promoter region, respectively.

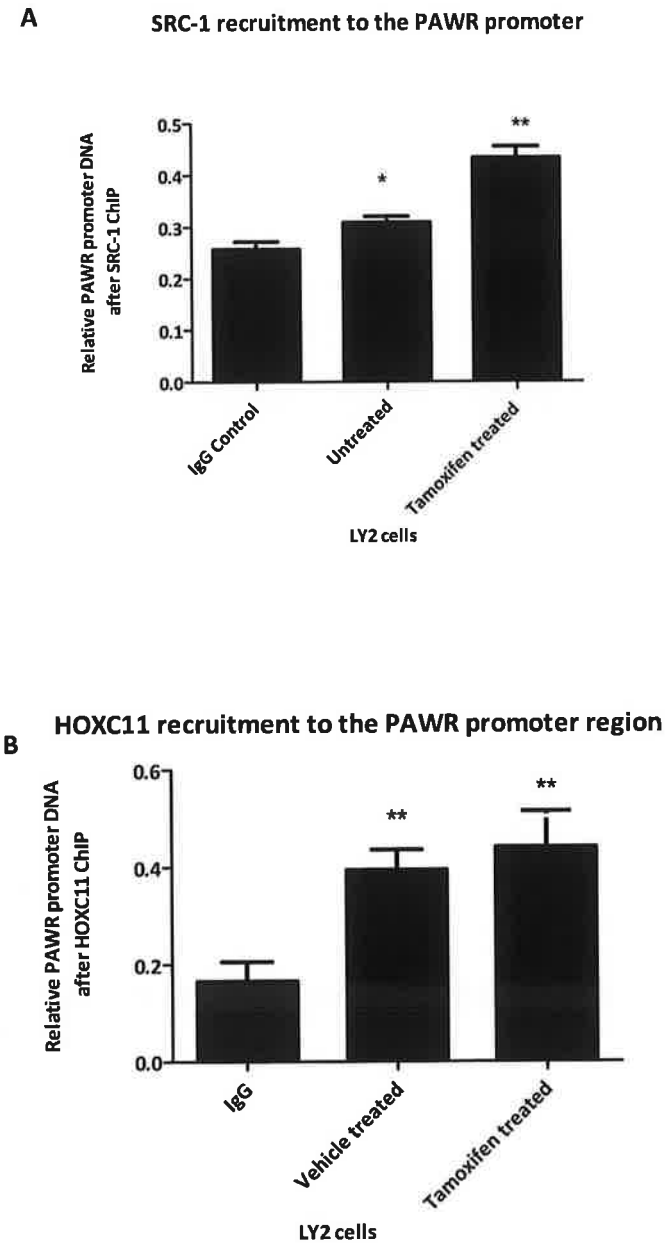


Figure 4.5. Chromatin Immunoprecipitation confirms recruitment of SRC-1 and HOXC11 to the PAWR promoter in LY2 endocrine resistant cells. ChIP analysis of **(A)** SRC-1 and **(B)** HOXC11 recruitment to the PAWR promoter was conducted to confirm the ChIPseq data sets. Quantitative real time analysis was used to quantify recruitment of both SRC-1 (n=3) protein and HOXC11 protein (n=5) to the DNA and data is represented relative to the IgG negative control (*,$p<0.05$, **,$p<0.001$, relative to the IgG control). Each sample is normalised to the specific experimental replicates.

4.3.4. PAWR expression is lost in the *in vitro* model of endocrine resistant breast cancer

As previously discussed, PAWR is most commonly cited with regards to its activity in the prostate. There is relatively little research which has documented PAWR activity in breast cancer, it was therefore important to ascertain the expression levels of this protein in the *in vitro* model of endocrine resistance, which had been used for all of the previous molecular work. mRNA from MCF-7 and LY2 cells was taken under basal conditions and analysed for PAWR gene expression using quantitative real time PCR analysis (Figure 4.6A). The results of the mRNA analysis reported higher levels of PAWR in the endocrine sensitive cell lines as opposed to the endocrine resistant cells. Western blot analysis was conducted using the same cell lines in order to confirm the concomitant loss of PAWR expression as tumour cells develop resistance (Figure 4.6B). Again, the protein was extracted from the MCF-7 and LY2 cell lines under basal conditions and the resultant protein analysis supported the RNA data as PAWR protein expression was dramatically reduced in the LY2 cells compared to the MCF-7 cells.

Additional, data from the LY2 cell line looked at the transcriptional response of PAWR to Tamoxifen treatment (Figure 4.4C). Following a period of steroid depletion, the cells were treated for either 1 hour or 4 hours with 10^{-7}M 4-OHT. The LY2 cells responded clearly to the effects of Tamoxifen as PAWR mRNA expression decreased in a manner that inversely correlated to the exposure of Tamoxifen.

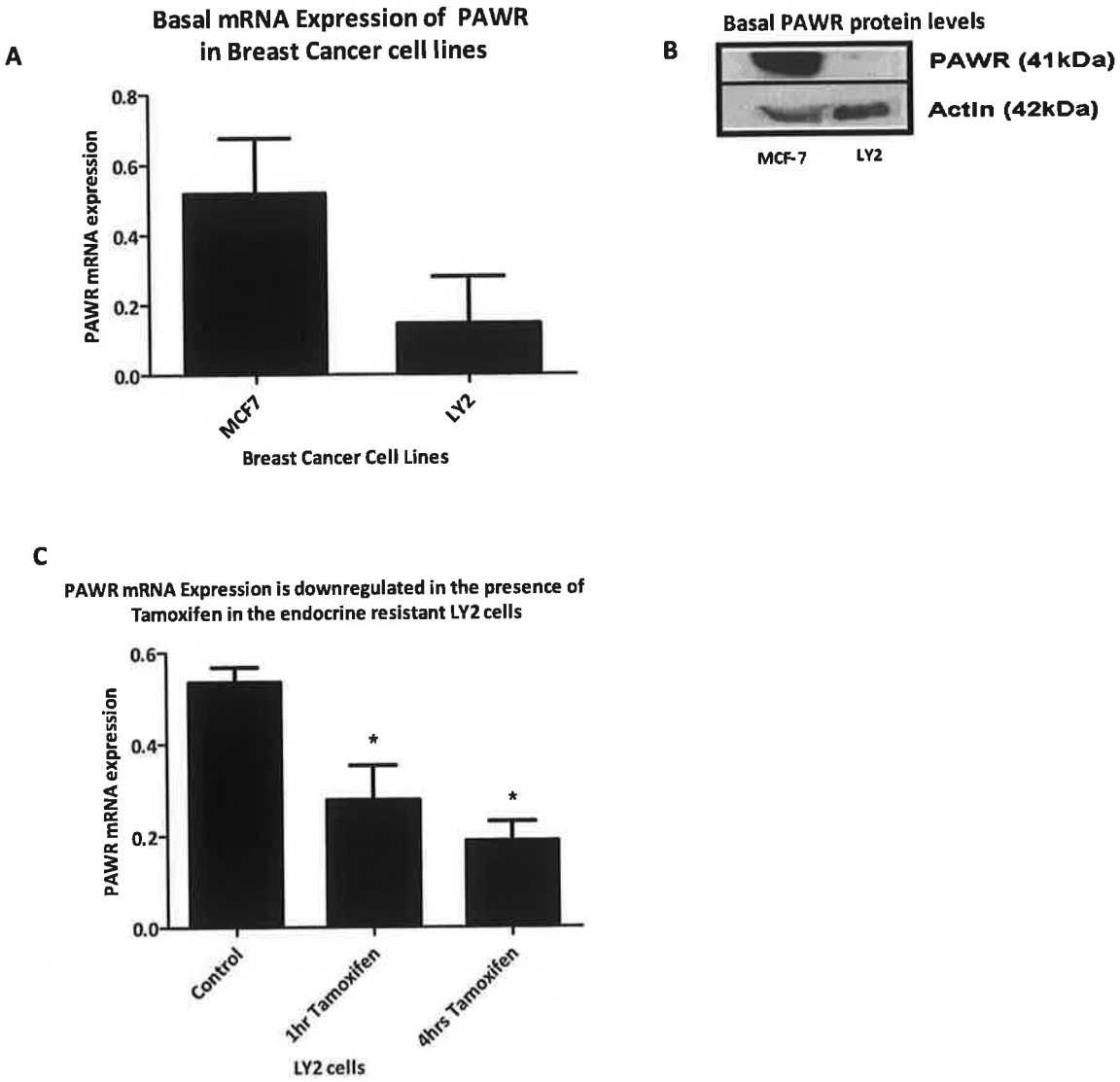


Figure 4.6. PAWR expression is downregulated in endocrine resistant breast cancer. (A) Quantitative real time PCR analysis of PAWR mRNA expression identified reduced levels of PAWR in the endocrine resistant LY2 cells compared to the endocrine sensitive MCF-7 cells (n=2). **(B)** MCF-7 and LY2 cells were lysed and run on a 10% polyacrylamide gel for PAWR. The gels were transferred and probed with a rabbit polyclonal antibody against PAWR. Western blot analysis showed a significant decrease in PAWR protein expression in the LY2 cells compared to the MCF-7 cells. The membrane was subsequently reprobed for β-actin which serves as a loading control for the experiment. **(C)** Quantitative real time PCR analysis of PAWR mRNA expression in LY2 cells which have been treated with vehicle, treated with 10^{-7}M Tamoxifen for 1 hour or treated with 10^{-7}M Tamoxifen for 4 hours (n=2,*, $p<0.05$ relative to the vehicle treated control) .

4.3.5. SRC-1 and HOXC11 negatively regulate PAWR in breast cancer

As published in McIlroy et al, 2010 and shown in Figure 3.4 and 3.11, the endocrine resistant LY2 cells have significantly upregulated SRC-1 and HOXC11 activity compared to the MCF-7 cells [62]. Furthermore, SRC-1 and HOXC11 only complex together in the LY2 cells thus these two cell lines represent an appropriate model in which to assess the regulation of PAWR transcriptional activity. Having confirmed that both SRC-1 and HOXC11 are recruited to the PAWR promoter, it was important to confirm their ability to regulate PAWR transcriptional activity. The ChIP experiments indicate that the SRC-1/HOXC11 can regulate PAWR but those experiments do not infer which direction the regulation occurs. Basal level expression of PAWR in the cell lines suggested that PAWR is negatively regulated by SRC-1/HOXC11 therefore transient transfection experiments were conducted in the same *in vitro* cellular model of resistance to confirm this hypothesised downregulation

cDNA was synthesised from SRC-1 and HOXC11 transient knockdown experiments in the LY2 cells to investigate the regulation of PAWR gene expression (Figure 4.7.). Quantitative real time PCR analysis was used to determine any changes in PAWR mRNA expression in response to the transient knockdown of either SRC-1 or HOXC11. The results confirmed that just as with CD24, PAWR transcriptional activity was also negatively influenced by the activity of SRC-1 and HOXC11. In the absence of SRC-1, the resistant LY2 cells exhibited increased levels of PAWR mRNA expression (Figure 4.7A). This expression pattern was repeated in the LY2s in which HOXC11 had been transiently knocked out (Figure 4.7C).

Again, transient overexpression transfections were carried in the MCF-7 cells with separate overexpression plasmids for both SRC-1 and HOXC11. The synthesised cDNA was analysed for PAWR gene expression and PAWR mRNA expression was shown to be markedly reduced in the overexpression samples. The opposing transfection experiments complemented each other well with all of the results supporting the negative regulation of PAWR by both SRC-1 and HOXC11 (Figure 4.7B, Figure 4.7D).

Additional evidence for PAWR as a downstream target of SRC-1 corepressor activity was obtained using the *SRC-1*$^{-/-}$ *MMTV-PyMT* mouse model from Prof Jianming Xu's lab. Quantitative real time PCR analysis of four wildtype tumours and four SRC-1 knockout tumours reported an increase in PAWR mRNA expression in the tumours that had no SRC-1 expressed (Figure 4.8A). This observation was further confirmed using protein lysate from the same tumours which showed that PAWR protein expression increased in the tumour cells that had no SRC-1 activity (Figure 4.8B). Altogether, the evidence suggests that PAWR

expression is lost in endocrine resistant breast cancer and that this loss is largely due to the negative influence of SRC-1 and HOXC11.

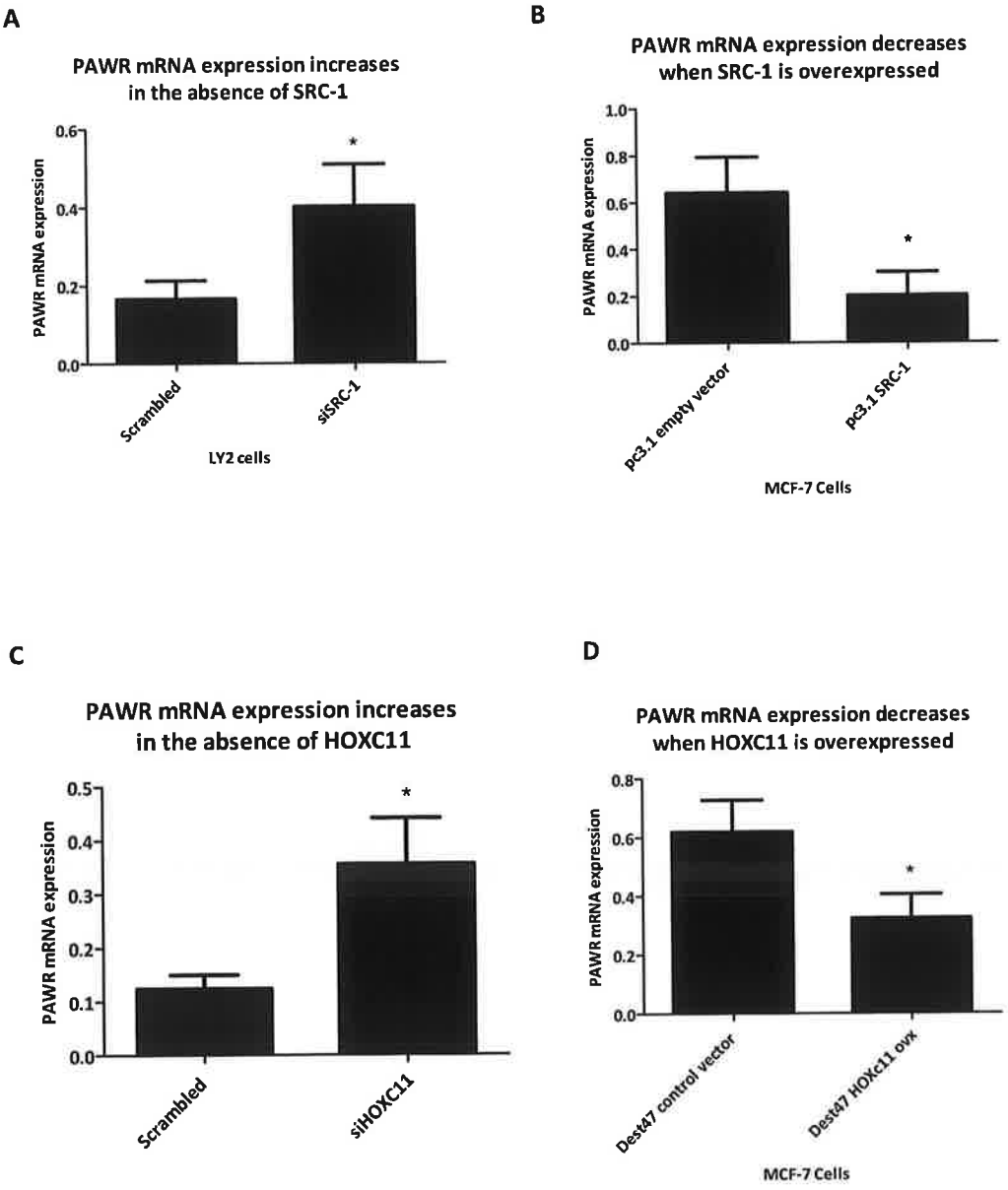
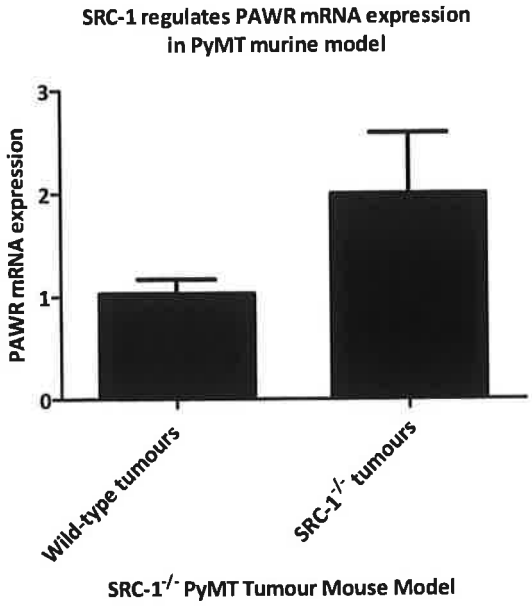


Figure 4.7. SRC-1 and HOXC11 negatively regulate PAWR transcription *in vitro*. Quantitative real time PCR analysis of PAWR mRNA expression in transiently transfected MCF-7 and LY2 cells. SRC-1 negatively regulates PAWR mRNA expression when it is **(A)** knocked out in LY2 cells (n=3) and when it is **(B)** overexpressed in MCF-7 cells (n=3). HOXC11 also negatively regulates PAWR mRNA expression when it is **(C)** knocked out in LY2 cells (n=3, $p<0.05$) and when it is **(D)** overexpressed in MCF-7 cells (n=4). *ovx: overexpression*

A



B

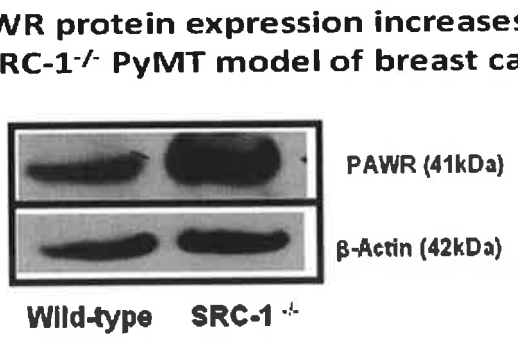


Figure 4.8. SRC-1 negatively regulates PAWR expression *in vivo*. (A) RNA was isolated from four *wildtype MMTV-PyMT* mammary gland tumours and four *SRC-1*$^{-/-}$*;MMTV-PyMT* mammary gland tumours. cDNA synthesis was performed and quantitative real time PCR analysis of PAWR mRNA expression was assessed in both sets of tumours. PAWR mRNA expression was increased in the *SRC-1*$^{-/-}$*;MMTV-PyMT* tumours in which the SRC-1 gene has been genetically silenced compared to the *wildtype MMTV-PyMT* tumours. **(B)** Cell lysates of wildtype and SRC-1 knockout tumours were ran on a 10% polyacrylamide gel, transferred and probed with a rabbit polyclonal antibody against PAWR. Western blot analysis was consistent with the PCR result and showed an increase of PAWR protein expression in the SRC-1 knockout tumour samples where SRC-1 activity had been ablated.

4.4. Discussion

The purpose of conducting ChIPseq experiments on both SRC-1 and HOXC11 was so that the results could be combined and common downstream target genes could be identified in an unbiased manner. The results of the HOXC11 ChIPseq experiment yielded far fewer peaks than the SRC-1 ChIPseq experiment (Figure 4.1A). This was to be expected given that SRC-1 is a much more dominant force in breast cancer than HOXC11. Furthermore, with the application of appropriate statistical cut-offs, the activity of HOXC11 in the resistant LY2 cells was almost completely driven by Tamoxifen treatment, whereby there was 1 peak in the untreated sample compared to 1,517 in the Tamoxifen treated sample. Consideration of this result in comparison to the SRC-1 ChIPseq peaks (ie 289 untreated peaks v 20486 Tamoxifen treated peaks; final ratio of 1:71), it is possible that HOXC11 activity in a resistant breast cancer context is far more dependent on the administration of Tamoxifen treatment than that of SRC-1. This was further evident in the images of the data using the UCSC genome browser (Figure 4.4B). Despite the smaller number of peaks obtained from the HOXC11 ChIPseq experiment, there was a final 54 overlapping peaks between the two data sets, of which 32 associated with protein coding genes (Table 4.1).

Since the SRC-1/HOXC11 complex appears to be capable of target gene activation and repression, ie S100β and CD24 respectively, the 32 new potential targets genes have potential to be regulated in either direction. Interestingly, PAWR, with its high levels of statistical significance was also a likely target of SRC-1 transcriptional corepression. Given that there has been no acknowledgement or investigation to date of SRC-1's potential to function as a corepressor in breast cancer, the appearance of PAWR, a tumour suppressor candidate gene, at the top of a list of potential SRC-1 targets was somewhat of a validation for the hypothesis of this work. Using the same approach as had been employed in the confirmation of CD24 as a genuine transcriptional target of SRC-1 and HOXC11 activity (Figure 3.14), the negative regulation of PAWR was also confirmed in both the *in vitro* cell model (Figure 4.7) and in the SRC-$1^{-/-}$ transgenic mouse model (Figure 4.8). Negative regulation of PAWR was confirmed at the transcriptional level in the *in vitro* model and the *in vivo* model. In addition, regulation was also observed at the protein level using the *in vivo* tumour cell lysates.

Previous research efforts with the PAWR protein have repeatedly associated it with endocrine related cancers such as prostate, endometrium and breast. These observations contributed strongly to the selection of PAWR as a possible target gene of SRC-1/HOXC11 activity and its subsequent investigation in this work. SRC-1 has associations with numerous

endocrine related cancers given its original function as a steroid receptor coregulator therefore in its efforts to progress a tumour, it is likely that SRC-1 would attempt to suppress a gene such as PAWR, whose functions is to tightly regulate apoptotic control over the tumour cell population. If PAWR activity is suppressed, then the cellular environment is increasingly susceptible to deregulation and hyper proliferation of inappropriate cell types. PAWR also has reputed associations with developmental biology which could explain its new found status as a target gene of HOXC11. An example of such is a study from 2009, which described the interaction between PAWR and the homeobox transcription factor, PITX2 (Pituitary homeobox transcription factor 2), in ocular cells. PITX2 expression is found during ocular development however aberrant expression of PITX2 has associations with an increased risk of glaucoma. PAWR binds PITX2 and functions to negatively regulate its transcriptional activity thus ensuring the normal development of the anterior segment of the eye [248].

The aim of this chapter was to identify additional target genes of the SRC-1/HOXC11 complex in endocrine resistant breast cancer. Although, PAWR has been validated as a target of negative regulation in this model and the general tumourigenic consequences associated with its loss make it an obvious target for SRC-1 and HOXC11, its suppression in this context may also bear a highly specific relevance to the development of resistance. The existence of tumour stem cells has been frequently addressed with regards to the development of endocrine resistance [249] [250], as both associate with a highly invasive and metastatic/migratory cancer phenotype. An important study from 2004, reported a PAWR-related mechanism by which to eliminate tumour stem cells [251]. In differentiating embryonic stem (ES) cells, those cells which expressed Oct-4 and PAWR, were induced to apoptose. Oct4 is a marker for pluripotency and also of tumour stem cells, it has also been associated with breast cancer. In the absence of PAWR, the Oct4 expressing stem cells survived and progressed to formed teratomas instead of differentiating into the normal tissue hierarchy. In conclusion, if PAWR has the capacity to kill any burgeoning stem cell populations within a tumour, its retained expression could have a genuinely positive effect on breast cancer prognosis. Moreover as the ChIPseq results attest, its targeted suppression could also have a genuinely positive effect on tumour progression.

5. Mechanisms of SRC-1 corepressor activity and the clinical consequences associated with SRC-1 targeted gene silencing

5.1. Introduction

5.1.1. DNA methylation

DNA promoter hypermethylation is one of the main epigenetic mechanisms of gene silencing and involves the addition of a methyl group to the carbon-5 position of the cytosine residue (5MeC) on the DNA. DNA methylation usually occurs at CpG islands, a CG rich region, upstream of the promoter region. The letter "p" here signifies that the C and G are connected by a phosphodiester bond. CpG islands are defined as sequences greater than 500bp in length, and have a GC content greater than 55%. Also approximately 60% of CpG islands are located in the promoter regions of protein coding genes [252]. In humans, DNA methylation is carried out by a group of enzymes called DNA methyltransferases (DNMTs). These enzymes not only determine the DNA methylation patterns during the early development, but are also responsible for copying these patterns to the strands generated from DNA replication. DNMT1 is the major player in maintenance of methylation in the genome. DNMT1 is often located at the replication fork during cellular division, allowing the copying of existing methylation onto the newly synthesized strand of DNA, thereby maintaining methylation patterns. In contrast, DNMT3A and DNMT3B act as *de novo* methyltransferases which remain active on unmethylated DNA and establish methylation patterns during early stages of development [253].

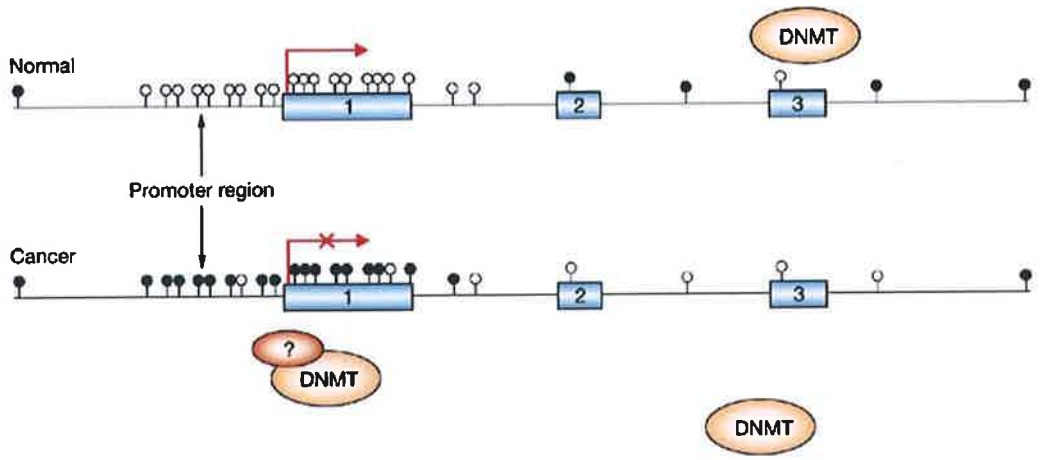


Figure 5.1. DNA methylation in normal and cancer cells. Schematic representation of DNA methylation of CpG sites (white circles) in a gene promoter region. Hypermethylation of CpG sites (black circles) results in disassociation of DNA methyltransferase (DNMT) and transcriptional silencing of the target gene. In normal tissues, CpG islands are mostly unmethylated. *Adapted from Baylin, 2005.*

In cancer cells, DNA methylation within gene promoter regions serves to turn off critical genes that could otherwise suppress tumourigenesis. Tumour suppressor genes and those encoding cell adhesion molecules and growth-regulatory proteins are often silenced in malignancies by DNA hypermethylation. It is believed that epigenetic gene inactivation is at least as common as, if not more frequent than, mutational events in the development of cancer. The large number of genes found to undergo hypermethylation in various malignancies suggests a role for epigenetic changes in the initiation and/or progression of cancer [254].

5.1.2. DNA methylation as a potential means of SRC-1 driven transcriptional repression

SRC-1 mediates the coactivation of transcriptional processes by recruitment of transcriptional machinery and specific enzymes to the transcriptional start site. This work has shown that SRC-1 also has the capacity to mediate the corepression of similar transcriptional processes, however the mechanism for this remains unknown. Given SRC-1's functional domains, it is likely that that SRC-1 streamlines the assembly of the corepressor complex in a similar manner to its assembly of the coactivator complex however the co-coactivators such as p300 and CBP could potentially be replaced by co-corepressors such as HDACs and DNMTs (Figure 5.2).

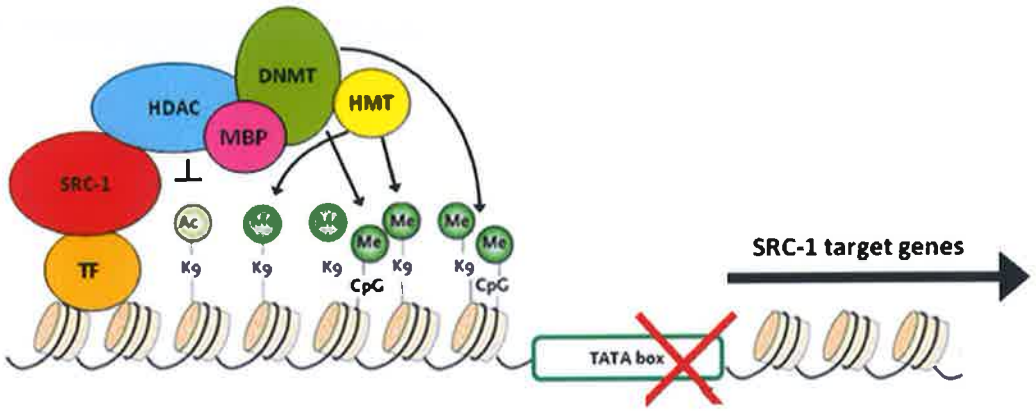


Figure 5.2. Proposed mechanism of SRC-1 corepressor activity in endocrine resistant breast cancer. SRC-1 forms steady state transcriptional complexes composed of numerous co-coactivators and general transcriptional machinery when it coactivates its many positively regulated target genes. It is therefore possible that in the opposing scenario where SRC-1 is functioning as a corepressor, it also dictates the assembly of necessary molecules such as histone deacetylases (HDACs), methyl-C binding proteins (MBPs) or histone methyltransferases (HMTs) in order to keep the chromatin in a condensed state and also specific DNA methyltransferases (DNMTs) required to maintain transcriptional quiescence of its specific target genes.

The most common nuclear corepressor (NCoR) has been associated with DNA methylation as a potential mechanism of transcriptional silencing. NCoR binds with a methyl CpG binding protein, Kaiso and together they bind to specific CpG rich sequences in a methylation dependent manner. In addition to Kaiso, NCoR also requires a functional histone deacetylase complex to mediate repression via DNA methylation [255].

Further literature searches also provided evidence that loss of CD24 and loss of PAWR have both been associated with promoter hypermethylation in differing physiological contexts. Aberrant DNA methylation was associated with altered transcription of CD24 in a common ocular surface disease known as pyterigium [256] Also unpublished data from Kaipparettu et al, saw evidence of CD24 promoter hypermethylation in breast cancer cell lines [220]. In endometrial cancer, promoter hypermethylation was shown to be responsible for PAWR silencing in both endometrial cancer cell lines and in primary tumour samples. Moreover, the frequency of PAWR promoter hypermethylation was 32% of all the analysed endometrioid endometrial carcinomas which ranks PAWR among the most frequently methylated genes in endometrial cancer [239].

5.1.3. Use of tissue microarray data to validate molecular observations

Due to the many limitations of immortalised cell lines, the ability to make clinical observations is becoming increasingly important as a means of validating molecular data. Tissue microarrays are a valuable and precious tool for clinical researchers as they enable miniaturisation of a large patient population and a high throughput approach to the analysis of intact tissues. In addition, to the acquisition and archiving of tumour tissue samples, TMA analysis is routinely accompanied by significant clinical history on each arrayed patient. Together this data can prove quite powerful if the patient cohort is larger enough to draw statistically significant conclusions. Kaplan-Meier estimates of survival are commonly employed to compute the survival over time, based on various clinical parameters. TMAs are also very useful for making clinical correlations with other established clinical pathological variables. Despite the sophisticated molecular classification systems employed to profile breast cancer subtypes, the situation in a routine clinical setting remains far more basic. In addition to scans and surgeries, clinicians rely on only a small number of markers which are identified via immunohistochemical means to diagnose a tumour and to strategise a treatment; they generally include ER, PR, HER2 and grade. If HER2 expression is borderline using immunohistochemistry, it is then reassessed by FISH (fluorescent in situ hybridisation) to identify any amplification of the HER2 gene. Tumour grade is routinely assessed by a trained pathologist using simple haematoxylin and eosin staining of the

tumour tissue to determine the architecture. In some breast cancer clinics, Ki67, the proliferative marker is also included to better gauge the nature of the tumour.

As outlined in section 1.1.3, receptor expression alone can be very informative for tumour cell characterisation. Correlation of CD24 and PAWR with ER, PR HER2 and grade as well as assessing their impact on survival in a large patient population was very important with regards to substantiating and supporting the observations which were made in the *in vitro* experimental model and the *in vivo* SRC1-/- transgenic mouse model.

5.2. Aims

- To investigate DNA promoter methylation as a possible mechanism of SRC-1-mediated corepressor activity.

- To validate the *in vitro* molecular observation of CD24 and PAWR regulation in a clinical patient population

- To investigate clinical correlations between CD24 and PAWR expression and other pathological markers of breast cancer.

5.3. Results

5.3.1. A potential mechanism for SRC-1 corepression of downstream target genes

This work has adopted a global approach to the investigation of SRC-1 action in the specific disease context of endocrine resistant breast cancer. The SRC-1 ChIPseq data identified the specific genomic regions where SRC-1 bound under the influence of Tamoxifen. In addition to this, the SRC-1 DNA microarray data identified the genes which were regulated in response to SRC-1 expression under the influence of Tamoxifen. Comparative analysis of the two datasets produced a list of 1,061 genes which were definitively downregulated by SRC-1 activity in endocrine resistance. To further specify how these 1,061 genes were being transcriptionally silenced, a SRC-1 specific global analysis of CpG methylation sites was also pursued in this same endocrine resistant context in order to identify a highly specific gene set whose expression was subject to SRC-1 directed suppression via DNA hypermethylation.

The methylated DNA immunoprecipitation (MeDIP) experiment was conducted in LY2 cells which were untransfected or had been transiently transfected with either scrambled control or with siSRC-1. The MeDIP samples were then arrayed to produce a list of genes which were specifically methylated in response to SRC-1 activity in the LY2 cells. Stringent analysis of the MeDIP microarray experiment clarified that CpG islands at gene promoter regions which had a peak score greater than 2 were significantly methylated. The results of the MeDIP microarray identified 1,297 methylated genes in the LY2 scrambled sample where SRC-1 is actively expressed. Further bioinformatic analysis showed that 964 genes of these 1,297 methylated genes were differentially methylated compared to the siSRC-1 LY2 sample where SRC-1 expression was ablated. These 964 methylated genes were therefore specifically methylated as a result of SRC-1 activity. In order to ascertain a definitive subset of genes which SRC-1 directly downregulated using DNA methylation, the 964 methylated genes were subsequently overlapped with the 1,061 SRC-1 downregulated genes which were identified from marrying the SRC-1 ChIPseq data with the SRC-1 DNA microarray data (Figure 3.7). The resulting overlap between the two subsets generated a list of 35 highly specific genes (Figure 5.3/Table 5.1). These genes represent a definitive cohort of target genes, specific to endocrine resistant disease, that are transcriptionally silenced via SRC-1 driven hypermethylation of their promoter regions (Figure 5.3).

Neither CD24 nor PAWR were identified by the MeDIP microarray using the stringent peak cut-off of 2, however SignalMap software was used to look at the methylation status of the specific CpG islands present in the promoter region of each gene. Although the methylation

Table 5.1. List of 35 genes which were directly downregulated by SRC-1 driven DNA methylation in endocrine resistant breast cancer

Chromosome	Gene Name	Peak Score
Chr. 1	CASP9	2.02
Chr. 1	MAN1A2	2.08
Chr. 1	KCNQ4	2.24
Chr. 1	POLR3C	2.79
Chr. 2	BCL2L11	2.61
Chr. 5	NR2F1	2.08
Chr. 6	GABBR1	2
Chr. 6	TAPBP	2.17
Chr. 8	TSNARE1	2.94
Chr. 8	EGR3	3.92
Chr. 11	ZDHHC13	2.23
Chr. 12	HRK	2.07
Chr. 13	ANKRD10	2.13
Chr. 15	TSSC1	2.15
Chr. 16	ANKRD11	2.3
Chr. 16	ATXN2L	2.31
Chr. 16	RAB11FIP3	2.44
Chr. 16	ZNF689	2.66
Chr. 17	MINK1	2.08
Chr. 17	ITGB4	2.22
Chr. 17	C1QL1	2.32
Chr. 17	NFE2L1	2.56
Chr. 17	ZMYND15	2.58
Chr. 17	CXCL16	2.58
Chr.19	BBC3	2.35
Chr. 19	STAP2	2.45
Chr. 19	PVRL2	2.47
Chr. 19	EPS8L1	2.5
Chr. 19	RPL36	2.5
Chr. 19	TRIP10	2.86
Chr. 19	MYO9B	4.82
Chr. 20	PARD6B	3.12
Chr. 21	RUNX1	2.22
Chr. 22	PLXNB2	2.6
Chr. 22	TBC1D22A	3.33

Genes are ranked in order of chromosomal number

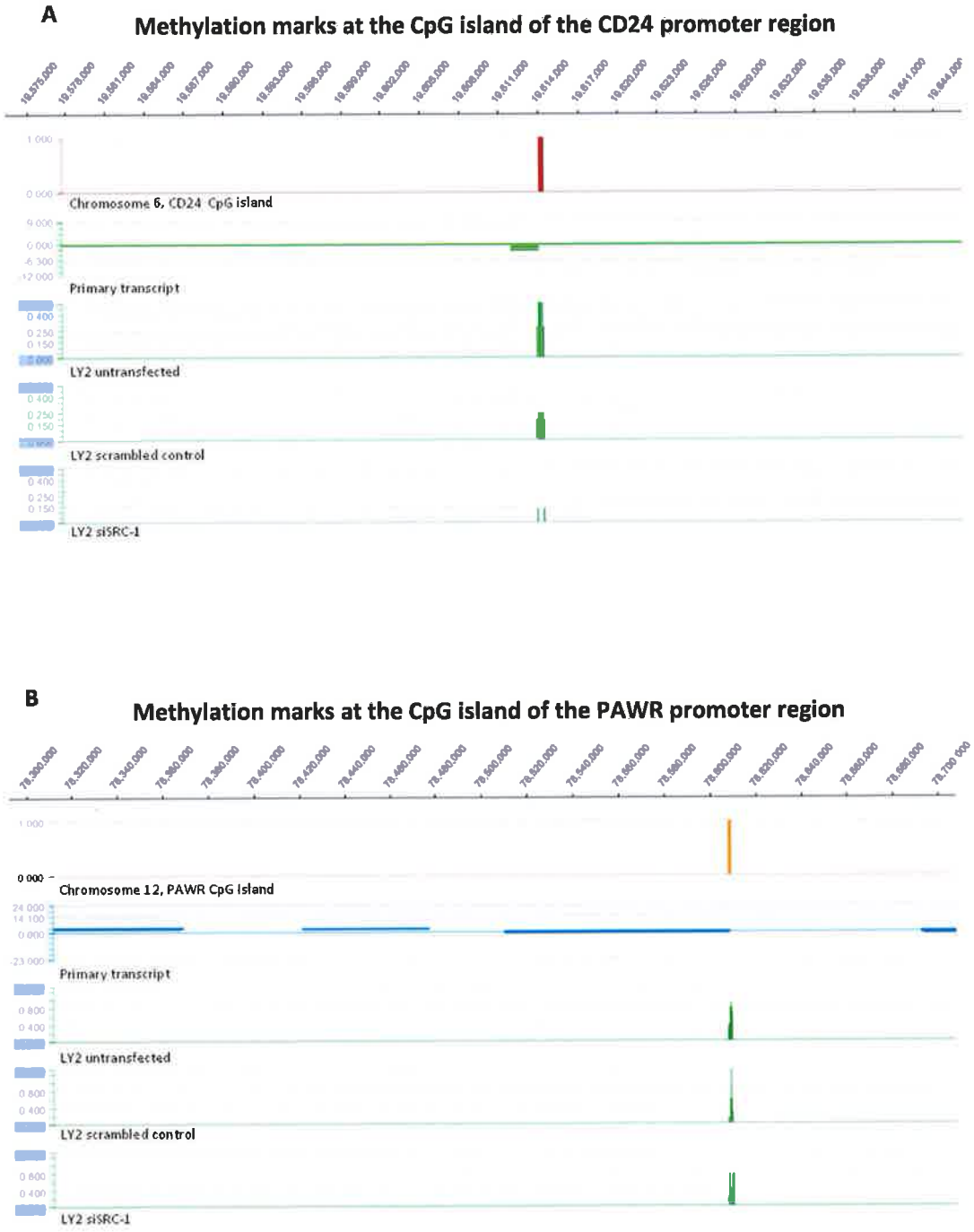


Figure 5.4. Methylation status of CD24 and PAWR in LY2 cells.

SignalMap images from MeDIP analysis for the upstream regions of **(A)** CD24 and **(B)** PAWR. The CpG islands for both genes are located close to the transcriptional start site of the primary transcript. Methylation marks at the CpG islands are shown in green for each experimental sample; LY2, LY2 scrambled control and LY2s transfected with siRNA against SRC-1.

5.3.2. Confirmation of SRC-1's functional relevance in a clinical patient population

A tissue microarray was previously constructed for this research group from primary tumours of 542 breast cancer patients who attended Beaumont hospital for treatment from 1996 – 2011. Each patient sample was represented by a minimum of four tissue cores on the TMA. The TMA was stained for SRC-1 protein expression using immunohistochemistry and the staining intensity was subsequently scored blindly by two observers.

A previous TMA had been constructed from a cohort of 560 patients who had attended St. Vincent's Hospital in Dublin. This TMA had also been stained by our group for SRC-1 in 2009 and the results were published by Redmond et al that same year [169]. The St Vincent's TMA reported that SRC-1 was predictive of poor disease free survival and showed increased predictive power when analysed in a Tamoxifen treated patient population. Importantly and in concurrence with observations of metastasis from the SRC-$1^{-/-}$ MMTV-PyMT mouse model, the St. Vincent's TMA also reported strong associations between SRC-1 positivity and nodal disease. The current Beaumont TMA was stained using the same SRC-1 antibody as the St. Vincent's TMA and the results confirmed and upheld the original observations with regards to the association of SRC-1 with poor disease free survival and nodal metastasis (Figure 5.5B / Table 5.2).

SRC-1 presents as a strong predictor of poor disease free survival in the overall Beaumont patient population. When the patient population is separated into untreated (Figure 5.5C) and Tamoxifen treated (Figure 5.5D) patient populations, SRC-1's ability to predict poor prognosis is specific only to the Tamoxifen treated cohort. This supports the proposal that SRC-1 has a definitive role to play in the development and mediation of endocrine resistance.

Using additional statistical analysis on the TMA data, further clinical correlations were made in the Beaumont TMA between SRC-1 and other indicators of disease progression. SRC-1 expression in breast cancer patients positively correlated with recurrence ($p<0.0001$) and with nodal disease ($p<0.05$), again reiterating the specific role of SRC-1 in resistance and metastasis (Table 5.2.).

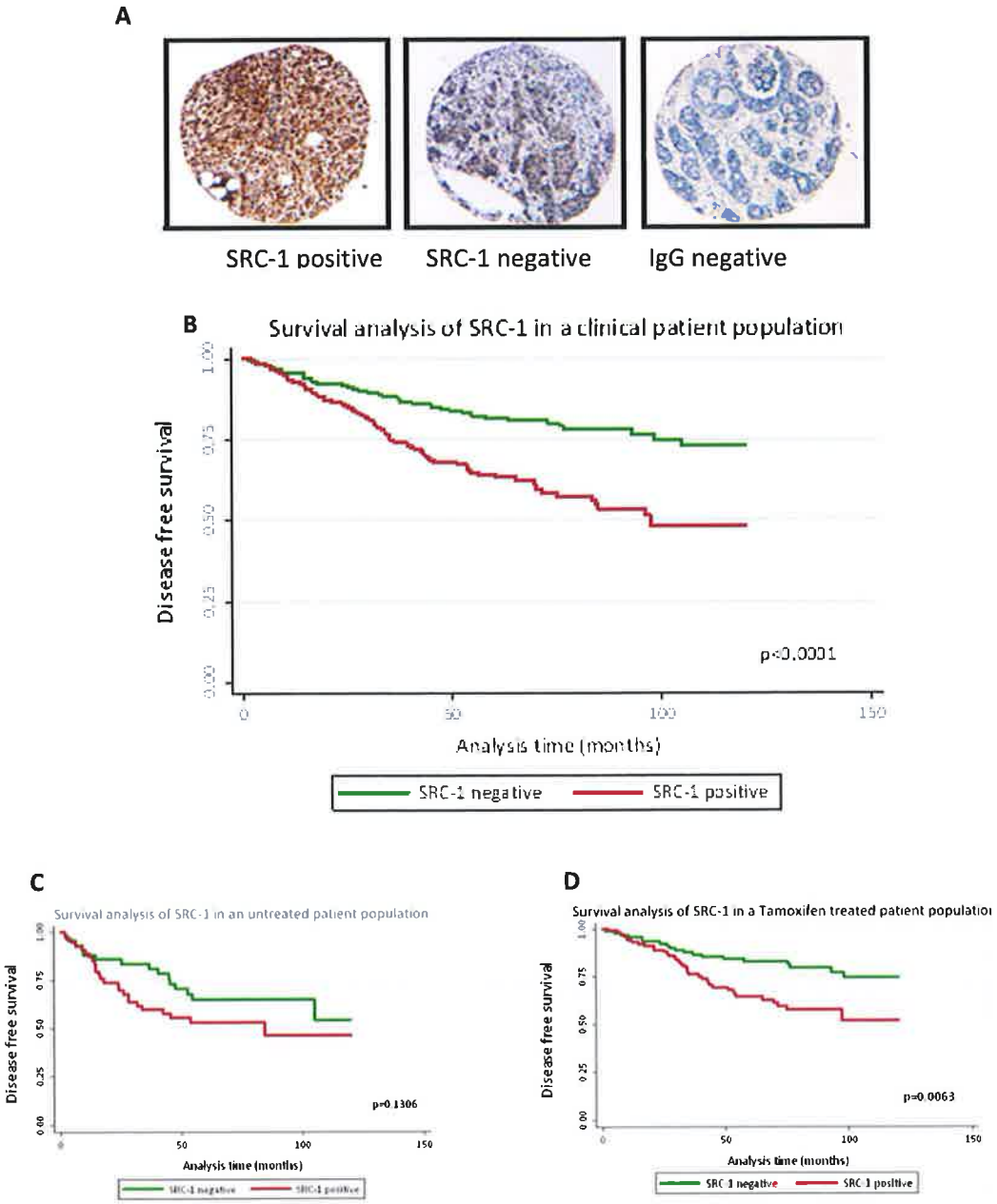


Figure 5.5. SRC-1 protein expression correlates specifically with poor prognosis in a Tamoxifen treated patient population.

(A) Representative tissue cores of PAWR staining in the TMA tissue cores. Kaplan Meier estimates of disease free survival for SRC-1 expression in the **(B)** overall (n=512), **(C)** untreated (n=101) and **(D)** tamoxifen treated (n=217) patient populations. SRC-1 positive primary tumours (red line) significantly associated with reduced disease free survival in the overall (p=<0.0001) and in the tamoxifen treated population (p=0.0063) but not significantly in the untreated population (p=0.1306). Survival analysis was reported using a 10 year period of patient follow-up.

Table 5.2. Positive associations of SRC-1 in the Beaumont breast cancer patient TMA using Fisher's exact test

	Recurrence	Nodal disease
SRC-1 +ve 55%	70%	60%
SRC-1 -ve 45%	30%	40%
SRC-1 associations	<0.0001	0.028
All patients (n=484)	31%	48%

5.3.3. CD24 expression correlates with a good disease prognosis in a clinical patient population

As outlined in Chapter 3, the MCF-7/LY2 *in vitro* model demonstrated that CD24 expression is lost as a tumour cell progresses from an endocrine sensitive to an endocrine resistant state. To verify whether these *in vitro* observations translated to a clinical patient population, the Beaumont TMA was stained for CD24 expression. Of the 352 primary breast cancer tumours stained with the CD24 antibody, 102 (29%) of the tumours were deemed positive for CD24 expression.

Overall survival analysis of CD24 in the patient population showed that CD24 expression associated with a good patient prognosis as determined by Kaplan Meier estimates of survival (Figure 5.6B). The inverse correlation observed between CD24 and poor disease free survival was highly significant ($p=0.0001$). Furthermore, additional estimates of survival reported that patients who were positive for CD24 and negative for SRC-1 had the best disease prognosis whilst patients who were negative for CD24 and positive for SRC-1 generally associated with a significantly worsened disease course (Figure 5.6C).

Importantly, statistical analysis supported the role of CD24 as a marker of well differentiated luminal breast cancer cells. CD24 expression was shown to positively correlate with ER expression ($p=0.002$) and also with the luminal A breast cancer patient subtype ($p<0.0001$) (Table 5.3). CD24 expression in breast cancer patients was also evaluated against other well established clinical pathological parameters. Specifically, CD24 expression exhibited a strong inverse correlation with SRC-1 expression which again is supportive of the molecular data that showed SRC-1 suppresses CD24 expression. CD24 also demonstrates significant inverse correlations with tumour grade ($p<0.05$), node positive disease ($p<0.01$) and disease recurrence ($p<0.0001$), all of which are common features of an aggressive cancer and a poor disease prognosis (Table 5.4).

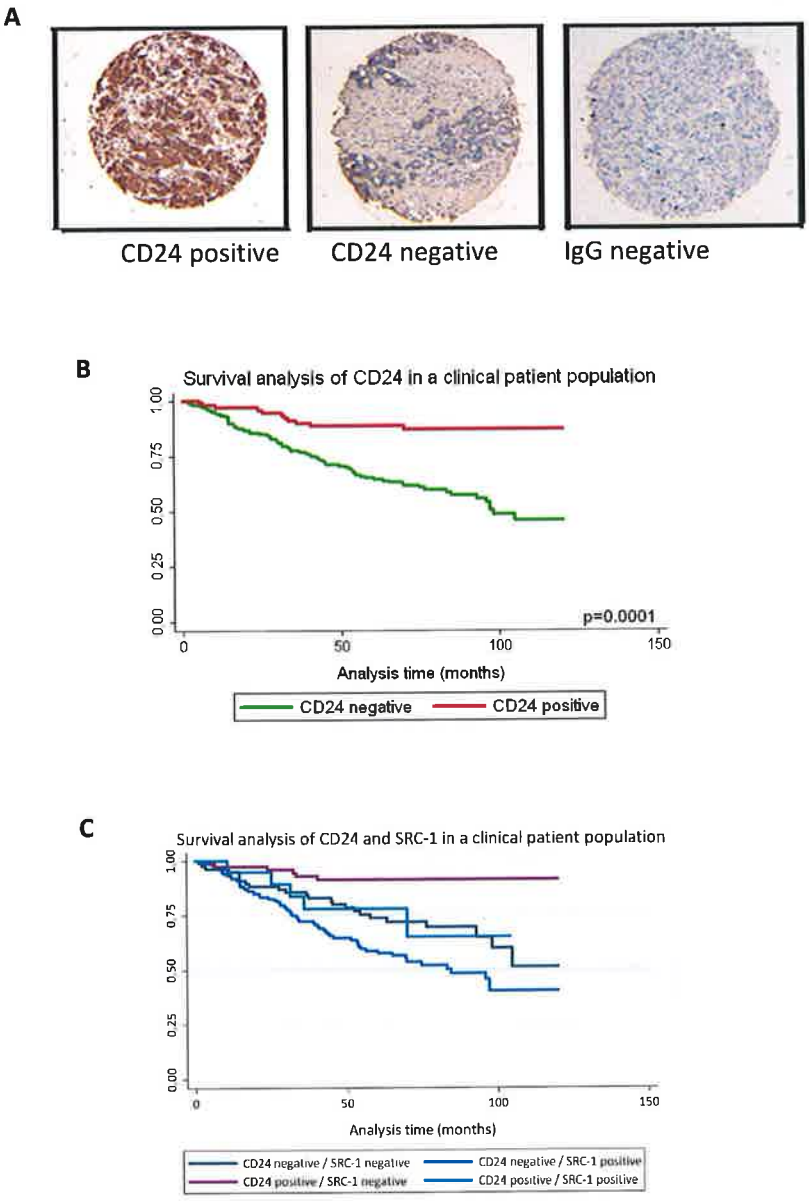


Figure 5.6. CD24 associates with good prognosis in breast cancer

(A) Representative tissue cores of CD24 staining in the TMA tissue cores. **(B)** Kaplan Meier estimates of disease free survival for CD24 expression in the overall patient population (n=352). CD24 positive primary tumours (red line) exhibit a clear inverse correlation with reduced disease free survival in the overall population ($p<0.0001$). **(C)** Kaplan Meier estimates of survival also analysed the strong inverse relationship between SRC-1 and CD24 expression in the breast cancer patient population. A survival function graph shows an inverse complex of SRC-1 positive and CD24 negative tumours significantly associates with poor rates of survival. Survival analysis was reported using a 10 year period of patient follow-up.

Table 5.3. Positive associations of CD24 in breast cancer patient TMA using Fisher's exact test

	Luminal A subtypes	ER expression	PAWR expression
CD24+ve 29%	36%	33%	47%
CD24-ve 71%	64%	67%	53%
CD24 associations	<0.0001	0.002	<0.0001
All patients (n=352)	55%	70%	22%

Table 5.4. Inverse associations of CD24 in breast cancer patient TMA using Fisher's exact test

	SRC-1 +ve	High grade	Nodal disease	Her2 expression	Recurrence
CD24+ve 29%	12%	24%	21%	23%	12%
CD24-ve 71%	88%	76%	79%	77%	88%
CD24 associations	*<0.0001	*0.049	*0.001	*0.095	*<0.0001
All patients (n=352)	52%	42%	47%	26%	32%

** signifies an inverse correlation*

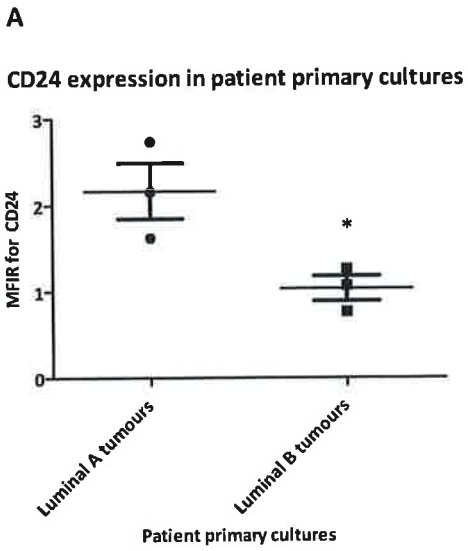
5.3.4. CD24 profiling in breast cancer primary cultures

Due to the generous participation of breast cancer patients undergoing treatment in Beaumont hospital and the dilligence of the surgical teams, primary cultures of tumour specimens were routinely received into the laboratory for research purposes. Since, the definition of certain tumour subtypes can be quite variable, the primary cultures received were classified into luminal A and luminal B tumour subtypes in accordance with the clinical guidelines of Beaumont hospital, The luminal A tumours were categorised as either ER and/or PR positive, HER2 negative and low grade. The luminal B tumours were categorised as ER and/or PR positive, HER2 positive or negative and high grade (Table 5.5). The tumour samples were processed and digested as outlined in section 2.1.8. A number of these primary cultures were made available for FACS analysis to determine their CD24 cell surface expression. Given the functional relationship of CD24 and CD44 with regards to profiling cellular populations for tumourigenicity and stem cell activity, the primary cultures were stained for CD24 and for CD44 in order to maximise the information yield from these rare samples. The primary cultures were analysed on the FACS Aria II flow cytometer in the same manner as the breast cancer cell lines had been. CD24 was reported in terms of percentage expression per total population and also mean fluoresence intensity (MFI). In line with the molecular observations, a ratio for the MFI of CD24 was calculated based upon the corresponding PE isotype control for each individual sample and the results showed that CD24 expression was significantly decreased in the luminal B tumours compared to the luminal A tumours (Figure 5.7A).

Representiative dot plots for the luminal A and luminal B subtypes were also shown for CD24 and CD44 (Figure 5.7B). The luminal A plot exhibits increased CD24 expression compared to the corresponding luminal B dot plot. The more aggressive luminal B tumour also has a distinct CD24-/CD44+ population which has been described in the literature as a clinically aggressive tumour cell population with possible cancer stem cell characteristics. Included with the dot plot is a representative profile of a patient who recurred on Tamoxifen therapy. Characterisation of this resistant tumour with CD24 and CD44 also produced a significant 'cancer stem cell' subpopulation and moreover had very little CD24 expression in the population as a whole. Ultimately , FACS analysis of CD24 expression in primary cultures concurred with the *in vitro* molecular data and the clinical TMA patient data and is supportive of an inverse relationship between CD24 and tumour aggression.

Table 5.5. Breast cancer primary cultures used for FACS analysis of CD24 expression

Primary tumour	ER	PR	HER2	Tumour Grade
Luminal A1	+	+	+	IDC/1
Luminal A2	+	+	+	ILC/1
Luminal A3	+	+	-	IDC/1
Luminal B1	+	+	+	IDC/3
Luminal B2	+	+	++	IDC/2
Luminal B3	+	-	+	ILC/2



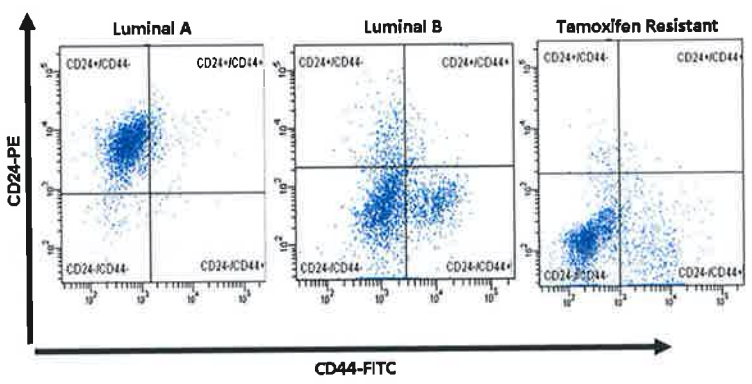


Figure 5.7. CD24 expression in breast cancer primary cultures

(A) FACS analysis of CD24 expression in luminal A tumours (n=3) and luminal B tumours (n=3). The MFIR of CD24 staining was calculated for both tumour cohorts and the results show that there was a significant loss of CD24 expression in the high grade luminal B tumours (*, $p<0.05$). **(B)** Representative dot plots of CD24 and CD44 staining of luminal A and luminal B breast cancer subtypes which also show a concomitant loss of CD24 in the luminal B tumours. The luminal B tumour has an increased CD24-/CD44+ subpopulation, the definition of which is commonly used to characterise the presence of cancer stem cell activity. The third dot plot profiles a Tamoxifen resistant tumour. As predicted, this tumour exhibits little CD24 expression and has a significant CD24-/CD44+ subpopulation.

5.3.5. PAWR expression correlates with a good disease prognosis in a clinical patient population

The tumour suppressor gene, PAWR was identified as a second target of SRC-1/HOXC11 corepression in the endocrine resistant breast cancer model. Just as had been done with CD24, PAWR expression was investigated in the TMA population in order to validate the molecular data obtained from the *in vitro* model. A total of 382 patient samples from the Beaumont TMA were stained using a rabbit polyclonal antibody against PAWR. Scoring of the tissue samples was conducted in the same manner as SRC-1 and CD24.

24% of the stained samples were recorded as positive for PAWR expression. Associated survival analysis of this 24 % in the overall population reported an inverse relationship between PAWR and poor disease free survival with a statistical significance of $p<0.01$ (Figure 5.8B).

In support of the survival analysis, PAWR expression also positively correlated with the luminal A tumour subtypes that had been identified within the TMA patient population (Table 5.6). Additional clinical parameters were also correlated against the PAWR expression data in order to construct a comprehensive profile of PAWR in a clinical setting. In the same pattern as CD24, PAWR demonstrated inverse associations to varying degrees of significance with high tumour grade, nodal disease, HER2 expression and recurrence (Table 5.7). Importantly, PAWR also showed a highly significant negative correlation with SRC-1 expression, this observation promotes the hypothesised negative regulation of PAWR by SRC-1. The TMA data strongly supports the evidence the original *in vitro* observation for SRC-1 mediated regulation of PAWR.

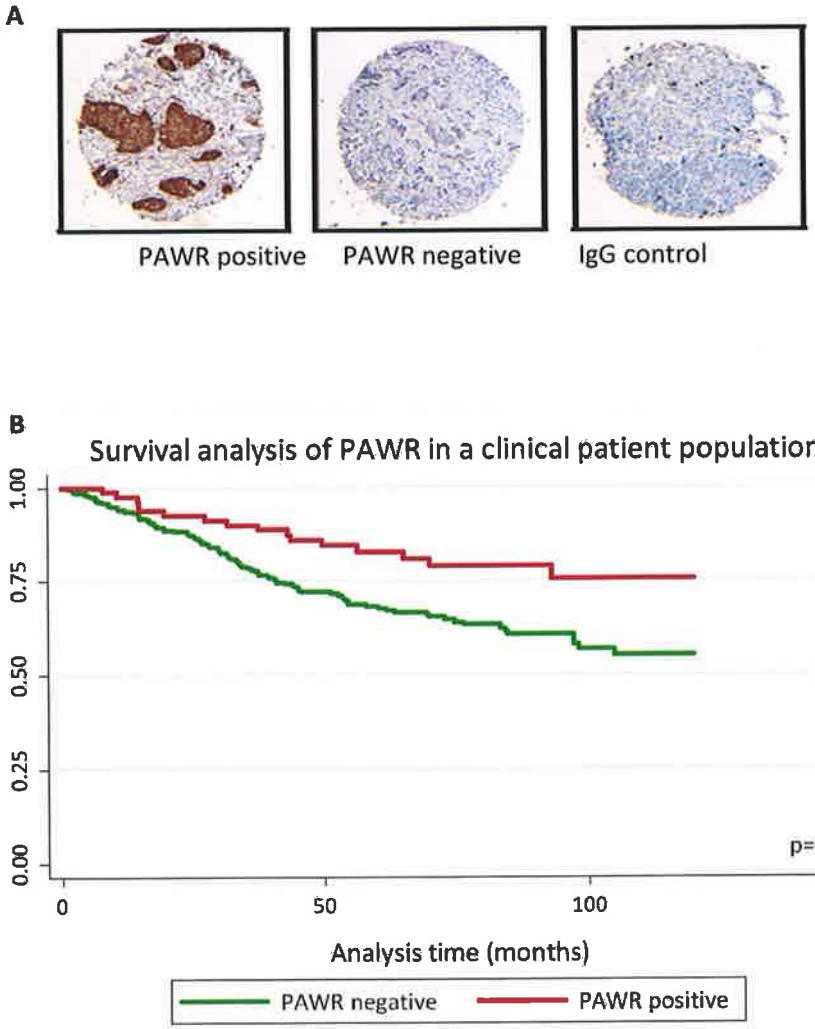


Figure 5.8. PAWR expression in breast cancer patient population

(A) Representative tissue cores of PAWR staining in the TMA tissue cores. (B) Kaplan Meier estimate of survival for PAWR expression in the overall TMA population. PAWR positive primary tumours (red line) exhibit a clear inverse correlation with reduced disease free survival in the overall population ($p<0.01$)

Table 5.6. Positive associations of PAWR in breast cancer patient TMA using Fisher's exact test

	Luminal A subtypes	ER expression
PAWR +ve 24%	33%	28%
PAWR-ve 76%	67%	72%
PAWR associations	<0.0001	0.011
All patients (n=382)	53%	69%

Table 5.7. Inverse associations of PAWR in breast cancer patient TMA using Fisher's exact test

	SRC-1 +ve	High grade	Nodal disease	Her2 expression	Recurrence
PAWR +ve 24%	15%	17%	19%	17%	14%
PAWR-ve 76%	85%	83%	81%	83%	86%
PAWR associations	*<0.0001	*0.004	*0.031	*0.021	*0.001
All patients (n=382)	53%	40%	47%	26%	32%

** signifies an inverse correlation*

5.4. Discussion

Transcriptional regulation of gene expression is subject to dynamic and transient regulatory events and processes. Transcriptional silencing of specific genomic regions can occur by interference with the recruitment of the RNA polymerase enzyme to the promoter, by specific recruitment of histone deacetylases or histone methyltransferases which ensure the chromatin is kept in a closed conformation so that the transcriptional machinery is unable to access the promoter regions. DNA methylation is an alternative mechanism of epigenetic regulation and has been identified as an increasingly relevant feature of cancer in numerous tumour types. Recruitment of DNA methyltransferases to methylate local CpG islands and prevent any transcriptional activation of specific genes has a more marked impact on gene expression than the more transitory histone modifications [257].

The MeDIP microarray experiment has highlighted 35 genes which are definitive targets of SRC-1 driven promoter methylation thus confirming that SRC-1 activity does associate with this specific mechanism of epigenetic regulation to mediate its corepressor function. Analysis of SRC-1's structure and its functional activity as a coactivator has shown that it is a nucleating point of contact in the assembly of transcriptional complexes; it is thus possible that SRC-1 could orchestrate recruitment of DNMTs and other corepressor molecules in a similar fashion as a means of transcriptionally silencing target genes such as CD24 and PAWR (Figure 5.2).

Due to technical issues resulting from this initial MeDIP microarray experiment, the results, although promising can only be considered as preliminary. The MeDIP microarray experiment will be repeated in LY2 cells which have been stably transfected with an SRC-1 short hairpin knockdown as a stable knockdown of SRC-1 should ensure increased reproducibility of the experimental data. Furthermore, rather than use the microarray platform, the MeDIP DNA will undergo whole genome sequencing to confirm the SRC-1 directed sites of methylation in a less biased manner than the arrays. The resultant methylation sites can then be further confirmed in the lab using bisulphite sequencing and methylation specific PCR analysis. Global analysis such as this will further elucidate the mechanisms of SRC-1 action in endocrine resistance and could also present opportunities for DNA methyltransferase inhibitors as possible targeted therapies in endocrine resistant breast cancer.

In addition to investigating a possible mechanism of action for SRC-1 corepressor activity, significant emphasis was placed on utilising available clinical data and resources to confirm

the functional consequences of SRC-1 and its novel target genes in a large breast cancer patient cohort. As previously mentioned, the Beaumont hospital TMA upon which all this data was analysed, is the second TMA constructed by this research group. Clinical observations have already been published from the first TMA and numerous proteins of interest investigated using this valuable research tool [62, 113, 169, 172]. SRC-1 expression has been analysed in both TMAs and the results are highly correlative thus attesting to the robustness of the datasets, the classifications and the analysis and ultimately lends confidence to any additional TMA findings.

Analysis of the TMA dataset successfully confirmed the molecular findings of this research. The patient data confirmed the anti-tumourigenic / pro-survival nature of both target genes of interest, CD24 and PAWR. Expression of CD24 and PAWR were both shown to positively correlate with luminal A tumour subtypes at a high level of statistical significance (Table 5.3 / Table 5.6). This observation thus supports the proposed association of CD24 and PAWR with well differentiated, steroid dependent, non aggressive tumour types. Importantly, the basic hypothesis of this research; the inverse relationship between SRC-1 and both target genes also held true in the clinical patient population, and moreover, this relationship was statistically significant for both genes. In support of this, CD24 and PAWR negatively correlated with other markers of disease aggression which again further substantiates their association with good prognosis and a regulated cellular environment (Table 5.4 / Table 5.7).

This research was fortunate in that, access to patient samples extended beyond the TMA slides and database. Following any breast surgeries conducted in Beaumont hospital, and with the provision of a large enough sample, primary cultures were collected and allocated for research purposes. To further validate, the CD24 TMA expression data, CD24 cell surface expression was analysed on these samples via flow cytometry and subsequently compared in tumours which had been subtyped as luminal A or luminal B. Distinct differences were not always clear based on cell surface expression alone therefore a ratio of mean fluorescence intensity was used to better quantify the staining intensity of each sample. As expected and in support of previous clinical observations, the loss of CD24 associates significantly with the higher grade luminal B tumour types (Figure 5.7).

To conclude, the analysis of the two anti-tumourigenic markers CD24 and PAWR in the clinical patient populations successfully confirmed the *in vitro* molecular observations. Both markers upheld their specific association with the luminal A breast cancer subtype whilst their targeted suppression distinctly associated with SRC-1 and with a worsened disease

course in breast cancer patients. The clinical consequences observed with loss of CD24 and PAWR confirm the relevance of these markers to the future investigations of breast cancer progression.

6. General Discussion

Current literature on SRC-1 activity in breast cancer has shown that SRC-1 can manipulate EMT, macrophage recruitment, receptor crosstalk, migration and differentiation in its efforts to promote tumourigenesis [74]. However all of these associations have been made as a result of SRC-1 coactivated transcriptional reactions. The predominant observation of this research is that the long standing steroid receptor coactivator, SRC-1, is also capable of significant corepressor activity in Tamoxifen resistant breast cancer. Genome wide analysis of aberrant SRC-1 activity in breast cancer highlighted the substantial number of genes which were transcriptionally silenced as a direct result of SRC-1 signalling. Targeted transcriptional silencing in breast cancer has commonly been associated with the renowned corepressors, NCoR and SMRT, therefore this is the first evidence to suggest that corepressor activity may also be directly attributed to the SRC-1 protein.

SRC-1 upholds an integral position in many transcriptional reactions where it directs the coordinated recruitment and assembly of the transcriptional machinery to the DNA. The wide diversity of transcription factors and associated transcriptional reactions that SRC-1 can coregulate, is indicative of how adaptable and functionally capable the SRC-1 protein can be. This unique capacity puts SRC-1 above those genes which are not involved in genomic regulation as a master regulator of cellular functions [50]. As a "master regulator" SRC-1 displays a sophisticated approach with its capacity to direct opposing yet coordinated transcriptional regulation of various cellular populations in order to construct signalling networks for tumour progression. Transcriptional regulation by its very nature is a transient and dynamic process and SRC-1 appears to be a molecule which can respond rapidly to the transcriptional landscape by asserting its functionality between that of a coactivator and that of a corepressor.

Global analysis of SRC-1 activity in breast cancer cell lines clearly showed that Tamoxifen treatment significantly impacts upon the activity of SRC-1 by driving it onto the DNA at many functional binding sites (Figure 3.6). However, the determining factors which influence SRC-1 to function as a coactivator or as a corepressor remain unknown. SRC-1 is a target of many environmental signalling pathways, such as the steroidogenic, cytokine and growth factor pathways, which can post-translationally instruct SRC-1 to simultaneously function in diverse manners [74]. Depending on the selective pressure upstream of SRC-1 at any given time or the immediate transcriptional landscape in a particular cellular context, SRC-1 could orchestrate its transcriptional agenda in differing ways with differing transcription factors. In the case of resistant breast cancer, as a tumour cell progresses, the ERα/SRC-1 functional partnership begins to unravel and deregulate and growth factor pathways become

increasingly influential. The various growth factor signalling molecules can inappropriately activate SRC-1, whereby it forgoes its associations with ERα or other previous partnerships to seek out new, often non steroidal binding partners, which can be utilised to protect the tumour cell and assist its survival.

The developmental protein HOXC11 is one such transcription factor. The functional relationship between SRC-1 and HOXC11 is completely specific to a tumourigenic context. Under normal circumstances, HOXC11 is predominantly active at embryogenesis where it coordinates development and cell fate decisions therefore its expression in the adult breast is entirely unexpected. However, cancer and normal development have common traits, as both processes involve shifts between cell proliferation and differentiation. With regards to HOXC11, SRC-1 may pursue this functional partnership to mediate reactivation of dormant developmental programmes in order to redirect tumour cell lineages along a less differentiated path, so that a tumour cell will become much more susceptible to tumourigenic pressures. HOX proteins are also deemed to be master regulators and when aberrantly expressed in cancer, they can function with multiple growth factor pathways that contribute to hormone resistant breast cancer and also metastasis [177]. Furthermore, the HOX family of proteins have been shown to regulate gene expression by both activation and repression; this adaptability makes HOXC11 an attractive binding partner for SRC-1 targeted action.

An important purpose of this work was to identify novel target genes of SRC-1 activity in breast cancer. Interestingly, the two target genes which were subsequently identified were both targets of SRC-1 suppression and not activation and furthermore both genes were strikingly similar in their anti-tumourigenic profiles. In contrast to the usual SRC-1 coactivated target genes, CD24 and PAWR both presented as effectors of good prognosis and are genuine contributors toward an appropriately controlled cellular environment. PAWR, in particular, has a definitive role in selecting certain cellular populations for targeted apoptosis in order to develop and maintain normally functioning tissue compartments. The development of resistance and disease progression in a tumour cell population coincides with a loss of normal cellular processing and increasing deregulation of intricate control mechanisms. The results presented in this thesis, suggest that it is the targeted loss of genes such as CD24 and PAWR which make it easier for a tumour cell to succumb to subsequent pro-survival or pro metastatic signals. Once significant control processes such as regulation of apoptosis, polarity, adhesion or differentiation become decommissioned, any remaining normal cellular compartments become functionally disorientated and are far more

vulnerable to tumourigenic assaults. SRC-1 appears to orchestrate the downregulation of genes which associate with good prognosis whilst at the same time upregulating the genes which associate with tumour progression, it is this dual functionality which makes SRC-1 such a potent effector of reduced disease free survival in clinical patient populations (Figure 5.1).

Given the wide-reaching oncogenic effects of SRC-1, investigations have been made to elucidate its potential as an effective biological therapeutic target. SRC-1 is a target of the major signalling pathways in cancer; through it they mediate their transcriptional preferences for tumour proliferation and metastasis (Figure 6.1). Targeting SRC-1 for therapeutic intervention would essentially stunt the operation of numerous signalling mechanisms at one specific point. Up to now, the SRC proteins had been disregarded as therapeutic targets due to their large size, lack of structural conformity and the absence of any high affinity ligand-binding domains. However, recent proof of principle experiments have challenged this view and shown that small molecule inhibitors (SMIs) such as gossypol, a natural polyphenol, are capable of reducing both SRC-1 and AIB1 proteins in MCF-7 and other cancer cell lines. This ability of gossypol to reduce SRC-1 and AIB1 proteins is also maintained independently of driving forces such as estrogen or Tamoxifen. Moreover, the gossypol action is independent of ERα as it also degrades AIB1 in ERα-negative cancer cells. Treatment of cancer cells with gossypol also resensitized the cells to MEK, EGF and IGF pathway inhibitors. Of significant importance, this form of SMI treatment is preferentially toxic to cancerous cells and is relatively selective for SRC-1 and AIB1. As SRC-1 activity is a hallmark of advanced disease or resistant disease, where treatment options have become substantially limited, a therapy that could specifically target SRC-1 expression could have profound implications for patient prognosis.

Regardless of targeted therapies specifically against SRC-1, the identification of its downstream targets and continued investigation of its associated signalling mechanisms will be a valuable asset for future therapeutic endeavours. As cancer genomics and bioinformatic analysis, become commonplace tools for breast cancer research, it has become increasingly apparent how sophisticated a disease, breast cancer is. Already, these methods are being translated into practical clinical applications, in particular gene expression assays such as Mammaprint and Oncotype DX [258]. Assays such as these, provide information as to the predicted survival or therapeutic response of a given tumour using gene expression signatures associated with that tumour. Such assays can also predict the clinical benefit of adjuvant therapy such as chemotherapy based upon the

characterization of the tumour expression signature. As a result, treatment options for patients must also become more sophisticated and specifically more personalized in order to combat a diagnosis in the most efficient and accurate way possible. Cancer genomics will continue to impact on the future of breast cancer treatment and global analysis of master regulator gene families such as the SRC and HOX proteins will contribute significantly to understanding specific subtypes of breast cancer.

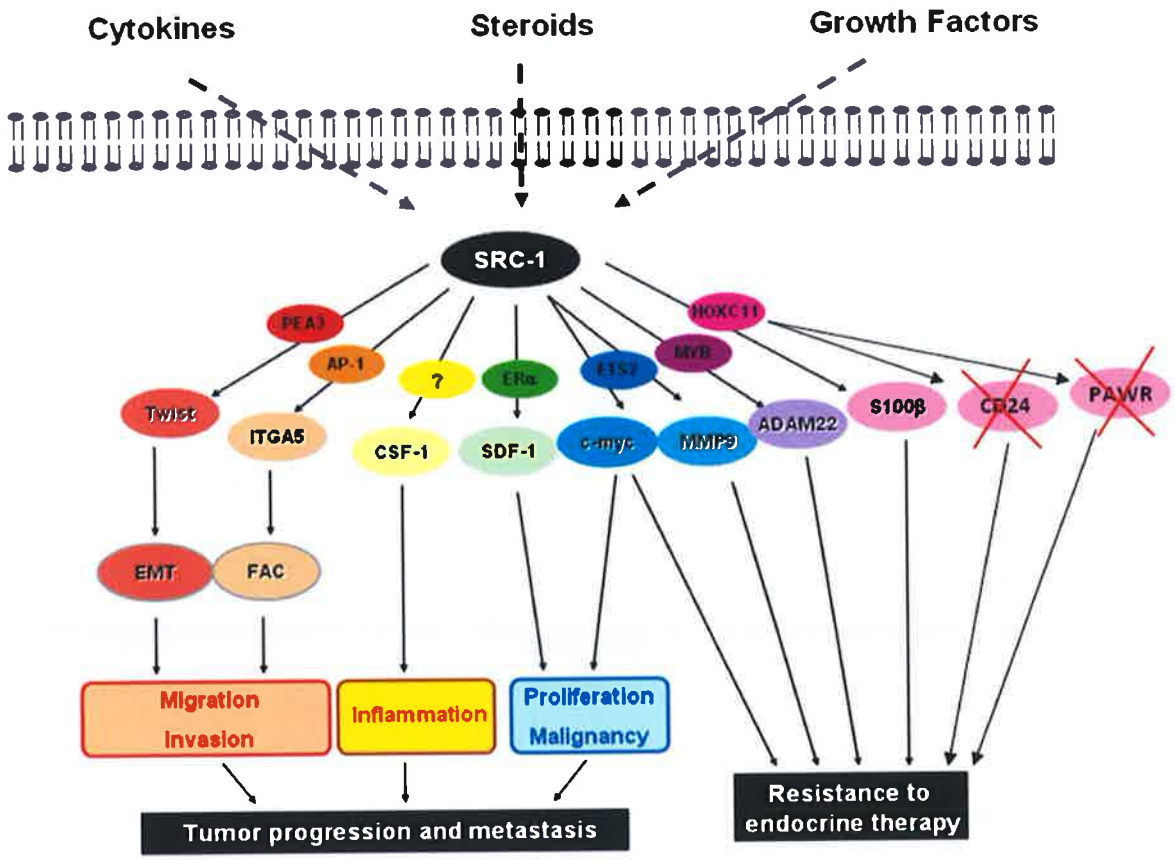


Figure 6.1 SRC-1 coregulator activity in breast cancer. SRC-1 is susceptible to activation by the cytokine, steroid or growth factor pathways. The schematic diagram illustrates the potential transcriptional binding partners of SRC-1 and their respective downstream effector targets. Under the influence of either steroid dependent or steroid independent signalling networks, SRC-1 will interact with different transcription factors to form varying complexes for transcriptional regulation of tumour cell processes that will promote an increasingly aggressive and resistant tumour phenotype. EMT: epithelial mesenchymal transition; FAC: focal adhesion complex. *Adapted from Walsh et al, 2012 [74].*

In conclusion, this research has confirmed the functional interaction of SRC-1 and the non steroidal transcription factor, HOXC11, in Tamoxifen resistant breast cancer. Furthermore it has identified a novel mechanism for SRC-1 regulation in breast cancer and extended the repertoire of SRC-1/HOXC11 target genes with the identification of CD24 and PAWR. Importantly, the clinical consequences of these SRC-1 interactions directly contribute to reduced disease free survival in Tamoxifen treated patient populations. In agreement with preceding investigations of SRC-1, this research supports the consensus that SRC-1 is a dominating and masterful force in the successful execution of diverse tumourigenic processes in resistant and metastatic breast cancer.

References

1. Lewis, M.T., *Faith, heresy and the cancer stem cell hypothesis.* Future Oncology, 2008. **4**(5): p. 585-589.
2. National Cancer Registry, C., Ireland, *Breast cancer incidence, mortality, treatment and survival in Ireland: 1994-2009.* 2012.
3. Ferlay, J., D.M. Parkin, and E. Steliarova-Foucher, *Estimates of cancer incidence and mortality in Europe in 2008.* European journal of cancer (Oxford, England : 1990), 2010. **46**(4): p. 765-781.
4. Doll, R., Peto, R., *The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today.* J Natl Cancer Inst, 1981. **66**(6): p. 1191-308.
5. Becker, N., *Breast Cancer Epidemiology.* Management of Breast Disease, Springer Press, 2010.
6. Blackwood, M.A. and B.L. Weber, *BRCA1 and BRCA2: from molecular genetics to clinical medicine.* Journal of Clinical Oncology, 1998. **16**(5): p. 1969-77.
7. Pharoah, P.D.P., et al., *Polygenes, Risk Prediction, and Targeted Prevention of Breast Cancer.* New England Journal of Medicine, 2008. **358**(26): p. 2796-2803.
8. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-752.
9. Sørlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets.* Proceedings of the National Academy of Sciences, 2003. **100**(14): p. 8418-8423.
10. Carey, L.A., et al., *Race, Breast Cancer Subtypes, and Survival in the Carolina Breast Cancer Study.* JAMA: The Journal of the American Medical Association, 2006. **295**(21): p. 2492-2502.
11. Hu, Z., et al., *The molecular portraits of breast tumors are conserved across microarray platforms.* BMC Genomics, 2006. **7**(1): p. 96.
12. Sorlie, T., et al., *Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms.* BMC Genomics, 2006. **7**(1): p. 127.
13. Rakha, E.A., et al., *Biologic and Clinical Characteristics of Breast Cancer With Single Hormone Receptor–Positive Phenotype.* Journal of Clinical Oncology, 2007. **25**(30): p. 4772-4778.
14. Bertucci, F., et al., *How different are luminal A and basal breast cancers?* International Journal of Cancer, 2009. **124**(6): p. 1338-1348.
15. Cheang, M.C.U., et al., *Ki67 Index, HER2 Status, and Prognosis of Patients With Luminal B Breast Cancer.* Journal of the National Cancer Institute, 2009. **101**(10): p. 736-750.
16. Sørlie, T., *Molecular portraits of breast cancer: tumour subtypes as distinct disease entities.* European Journal of Cancer, 2004. **40**(18): p. 2667-2675.
17. Bergamaschi, A., et al., *Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome.* The Journal of Pathology, 2008. **214**(3): p. 357-367.
18. Sotiriou, C., et al., *Breast cancer classification and prognosis based on gene expression profiles from a population-based study.* Proceedings of the National Academy of Sciences, 2003. **100**(18): p. 10393-10398.
19. Matros, E., et al., *BRCA1 promoter methylation in sporadic breast tumors: relationship to gene expression profiles.* Breast Cancer Research and Treatment, 2005. **91**(2): p. 179-186.
20. Turner, N.C., et al., *BRCA1 dysfunction in sporadic basal-like breast cancer.* Oncogene, 2006. **26**(14): p. 2126-2132.

21. Curtis, C., et al., *The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups.* Nature, 2012. **486**(7403): p. 346-352.
22. Osborne, C.K. and R. Schiff, *Mechanisms of Endocrine Resistance in Breast Cancer.* Annual Review of Medicine, 2011. **62**(1): p. 233-247.
23. Russo, J. and I.H. Russo, *Development of the human breast.* Maturitas, 2004. **49**(1): p. 2-15.
24. Visvader, J.E., *Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis.* Genes & Development, 2009. **23**(22): p. 2563-2577.
25. Hennighausen, L. and G.W. Robinson, *Signaling Pathways in Mammary Gland Development.* Developmental Cell, 2001. **1**(4): p. 467-475.
26. Taylor-Papadimitriou, J., et al., *Keratin expression in human mammary epithelial cells cultured from normal and malignant tissue: relation to in vivo phenotypes and influence of medium.* Journal of Cell Science, 1989. **94**(3): p. 403-413.
27. Mahendran, R.S., et al., *A new monoclonal antibody to a cell-surface antigen that distinguishes luminal epithelial and myoepithelial cells in the rat mammary gland.* Journal of Cell Science, 1989. **94**(3): p. 545-552.
28. Petersen, O.W., P.E. Høyer, and B. van Deurs, *Frequency and Distribution of Estrogen Receptor-positive Cells in Normal, Nonlactating Human Breast Tissue.* Cancer Research, 1987. **47**(21): p. 5748-5751.
29. Clarke, R.B., *Ovarian steroids and the human breast: Regulation of stem cells and cell proliferation.* Maturitas, 2006. **54**(4): p. 327-334.
30. Abd El-Rehim, D.M., et al., *Expression of luminal and basal cytokeratins in human breast carcinoma.* The Journal of Pathology, 2004. **203**(2): p. 661-671.
31. Goodman, M., *Basic Medical Endocrinology.* 2009: p. 257-275.
32. Engel, L.L., *The biosynthesis of estrogens* Cancer, 1957. **10**(4): p. 711-5.
33. Green, S., et al., *Human oestrogen receptor cDNA: sequence, expression and homology to v-erb-A.* Nature, 1986. **320**(6058): p. 134-139.
34. Greene, G., et al., *Sequence and expression of human estrogen receptor complementary DNA.* Science, 1986. **231**(4742): p. 1150-1154.
35. Chawla, A., et al., *Nuclear Receptors and Lipid Physiology: Opening the X-Files.* Science, 2001. **294**(5548): p. 1866-1870.
36. Weatherman, R.V., R.J. Fletterick, and T.S. Scanlan, *NUCLEAR-RECEPTOR LIGANDS AND LIGAND-BINDING DOMAINS.* Annual Review of Biochemistry, 1999. **68**(1): p. 559-581.
37. Kumar, V., et al., *Functional domains of the human estrogen receptor.* Cell, 1987. **51**(6): p. 941-951.
38. Nilsson , S. and J.-A. Gustafsson, *Estrogen receptor transcription and transactivation: Basic aspects of estrogen action.* Breast Cancer Res, 2000. **2**(5): p. 360 - 366.
39. Zilli, M., et al., *Molecular mechanisms of endocrine resistance and their implication in the therapy of breast cancer.* Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 2009. **1795**(1): p. 62-81.
40. Kuiper, G.G.J.M., et al., *Comparison of the Ligand Binding Specificity and Transcript Tissue Distribution of Estrogen Receptors α and β.* Endocrinology, 1997. **138**(3): p. 863-870.
41. McInerney, E.M., et al., *Transcription Activation by the Human Estrogen Receptor Subtypeβ (ERβ) Studied with ERβ and ERα Receptor Chimeras.* Endocrinology, 1998. **139**(11): p. 4513-4522.
42. Hall, J.M. and D.P. McDonnell, *The Estrogen Receptor β-Isoform (ERβ) of the Human Estrogen Receptor Modulates ERα Transcriptional Activity and Is a Key Regulator of the Cellular Response to Estrogens and Antiestrogens.* Endocrinology, 1999. **140**(12): p. 5566-5578.

43. Hennighausen, L. and G.W. Robinson, *Information networks in the mammary gland.* Nat Rev Mol Cell Biol, 2005. **6**(9): p. 715-725.
44. Förster, C., et al., *Involvement of estrogen receptor β in terminal differentiation of mammary gland epithelium.* Proceedings of the National Academy of Sciences, 2002. **99**(24): p. 15578-15583.
45. Klein-Hitpass, L., et al., *The progesterone receptor stimulates cell-free transcription by enhancing the formation of a stable preinitiation complex.* Cell, 1990. **60**(2): p. 247-257.
46. Kim, J. and M. Stallcup, *P160 Coactivators: Critical Mediators of Transcriptional Activation by Nuclear Receptors.* World Scientific Publishing Co. Pte. Ltd., Singapore., 2008.
47. Meyer, M.-E., et al., *Steroid hormone receptors compete for factors that mediate their enhancer function.* Cell, 1989. **57**(3): p. 433-442.
48. Onate , S., et al., *Sequence and characterization of a coactivator for the steroid hormone receptor superfamily.* Science, 1995. **24**(270): p. 1354-7.
49. Lonard, D.M. and B.W. O'Malley, *Nuclear Receptor Coregulators: Judges, Juries, and Executioners of Cellular Regulation.* Molecular Cell, 2007. **27**(5): p. 691-700.
50. O'Malley, B.W., *Masters of the genome.* Nat Rev Mol Cell Biol, 2010. **11**(5): p. 311-311.
51. Johnson, A.B. and B.W. O'Malley, *Steroid receptor coactivators 1, 2, and 3: Critical regulators of nuclear receptor activity and steroid receptor modulator (SRM)-based cancer therapy.* Molecular and Cellular Endocrinology, 2012. **348**(2): p. 430-9.
52. Cvoro, A., et al., *Distinct Roles of Unliganded and Liganded Estrogen Receptors in Transcriptional Repression.* Molecular cell, 2006. **21**(4): p. 555-564.
53. Berghagen, H., et al., *Corepressor SMRT Functions as a Coactivator for Thyroid Hormone Receptor T3Rα from a Negative Hormone Response Element.* Journal of Biological Chemistry, 2002. **277**(51): p. 49517-49522.
54. Rogatsky, I., et al., *Alternate surfaces of transcriptional coregulator GRIP1 function in different glucocorticoid receptor activation and repression contexts.* Proceedings of the National Academy of Sciences, 2002. **99**(26): p. 16701-16706.
55. Wang, J.-C., J.M. Stafford, and D.K. Granner, *SRC-1 and GRIP1 Coactivate Transcription with Hepatocyte Nuclear Factor 4.* Journal of Biological Chemistry, 1998. **273**(47): p. 30847-30850.
56. Zhu, Y., et al., *Cloning and identification of mouse steroid receptor coactivator-1 (mSRC-1), as a coactivator of peroxisome proliferator-activated receptor gamma.* Gene Expression, 1996. **6**(3): p. 185-95.
57. Ding, X.F., et al., *Nuclear Receptor-Binding Sites of Coactivators Glucocorticoid Receptor Interacting Protein 1 (GRIP1) and Steroid Receptor Coactivator 1 (SRC-1): Multiple Motifs with Different Binding Specificities.* Molecular Endocrinology, 1998. **12**(2): p. 302-313.
58. Lee, S.-K., et al., *Steroid Receptor Coactivator-1 Coactivates Activating Protein-1-mediated Transactivations through Interaction with the c-Jun and c-Fos Subunits.* Journal of Biological Chemistry, 1998. **273**(27): p. 16651-16654.
59. Kim, H.-J., J.H. Kim, and J.W. Lee, *Steroid Receptor Coactivator-1 Interacts with Serum Response Factor and Coactivates Serum Response Element-mediated Transactivations.* Journal of Biological Chemistry, 1998. **273**(44): p. 28564-28567.
60. Na, S.-Y., et al., *Steroid Receptor Coactivator-1 Interacts with the p50 Subunit and Coactivates Nuclear Factor κB-mediated Transactivations.* Journal of Biological Chemistry, 1998. **273**(18): p. 10831-10834.

61. Myers, E., et al., *Associations and Interactions between Ets-1 and Ets-2 and Coregulatory Proteins, SRC-1, AIB1, and NCoR in Breast Cancer.* Clinical Cancer Research, 2005. **11**(6): p. 2111-2122.
62. McIlroy, M., et al., *Interaction of Developmental Transcription Factor HOXC11 with Steroid Receptor Coactivator SRC-1 Mediates Resistance to Endocrine Therapy in Breast Cancer.* Cancer Research, 2010. **70**(4): p. 1585-1594.
63. Qin, L., et al., *Steroid Receptor Coactivator-1 Upregulates Integrin α5 Expression to Promote Breast Cancer Cell Adhesion and Migration.* Cancer Research, 2011. **71**(5): p. 1742-1751.
64. Qin, L., et al., *The Steroid Receptor Coactivator-1 Regulates Twist Expression and Promotes Breast Cancer Metastasis.* Cancer Research, 2009. **69**(9): p. 3819-3827.
65. Kim, J.H. and M.R. Stallcup, *P160 Coactivators: Critical Mediators of Transcriptional Activation by Nuclear Receptors.* World Scientific Publishing Co. Pte Ltd. Singapore, 2008.
66. Kim, J.H., H. Li, and M.R. Stallcup, *CoCoA, a Nuclear Receptor Coactivator which Acts through an N-Terminal Activation Domain of p160 Coactivators.* Molecular Cell, 2003. **12**(6): p. 1537-1549.
67. Zhang, H., et al., *Differential gene regulation by the SRC family of coactivators.* Genes & Development, 2004. **18**(14): p. 1753-1765.
68. Chang, C.-y., et al., *Dissection of the LXXLL Nuclear Receptor-Coactivator Interaction Motif Using Combinatorial Peptide Libraries: Discovery of Peptide Antagonists of Estrogen Receptors α and β.* Molecular and Cellular Biology, 1999. **19**(12): p. 8226-8239.
69. Heery, D.M., et al., *A signature motif in transcriptional co-activators mediates binding to nuclear receptors.* Nature, 1997. **387**(6634): p. 733-736.
70. Lanz, R.B., D.M. Lonard, and B.W. O'Malley, *Nuclear Receptor Coregulators in Human Diseases.* World Scientific Publishing Co. Pte Ltd. Singapore., 2008.
71. Coulthard, V.H., S. Matsuda, and D.M. Heery, *An Extended LXXLL Motif Sequence Determines the Nuclear Receptor Binding Specificity of TRAP220.* Journal of Biological Chemistry, 2003. **278**(13): p. 10942-10951.
72. Shiau, A.K., et al., *The Structural Basis of Estrogen Receptor/Coactivator Recognition and the Antagonism of This Interaction by Tamoxifen.* Cell, 1998. **95**(7): p. 927-937.
73. Bourguet, W., P. Germain, and H. Gronemeyer, *Nuclear receptor ligand-binding domains: three-dimensional structures, molecular interactions and pharmacological implications.* Trends in Pharmacological Sciences, 2000. **21**(10): p. 381-388.
74. Walsh, C., Qin, L., Tien, JCY., Young, LS., Xu, J. , *The Function of Steroid Receptor Coactivator-1 in Normal Tissues and Cancer.* . Int J Biol Sci 2012. **8**(4): p. 470-485.
75. Kraus, W.L., E.T. Manning, and J.T. Kadonaga, *Biochemical Analysis of Distinct Activation Functions in p300 That Enhance Transcription Initiation with Chromatin Templates.* Molecular and Cellular Biology, 1999. **19**(12): p. 8123-8135.
76. Chen, H., et al., *Nuclear Receptor Coactivator ACTR Is a Novel Histone Acetyltransferase and Forms a Multimeric Activation Complex with P/CAF and CBP/p300.* Cell, 1997. **90**(3): p. 569-580.
77. Spencer, T.E., et al., *Steroid receptor coactivator-1 is a histone acetyltransferase.* Nature, 1997. **389**(6647): p. 194-198.
78. Yang, X.-J., et al., *A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A.* Nature, 1996. **382**(6589): p. 319-324.
79. Voegel, J., et al., *TIF2, a 160 kDa transcriptional mediator for the ligand-dependent activation function AF-2 of nuclear receptors.* EMBO J, 1996. **15**(14): p. 3667-3675.
80. Onate, S.A., et al., *The Steroid Receptor Coactivator-1 Contains Multiple Receptor Interacting and Activation Domains That Cooperatively Enhance the Activation*

Function 1 (AF1) and AF2 Domains of Steroid Receptors. Journal of Biological Chemistry, 1998. **273**(20): p. 12101-12108.

81. Voegel, J.J., et al., *The coactivator TIF2 contains three nuclear receptor-binding motifs and mediates transactivation through CBP binding-dependent and -independent pathways.* EMBO J, 1998. **17**(2): p. 507-519.
82. Li, J., B.W. O'Malley, and J. Wong, *p300 Requires Its Histone Acetyltransferase Activity and SRC-1 Interaction Domain To Facilitate Thyroid Hormone Receptor Activation in Chromatin.* Molecular and Cellular Biology, 2000. **20**(6): p. 2031-2042.
83. Nakajima, T., et al., *RNA Helicase A Mediates Association of CBP with RNA Polymerase II.* Cell, 1997. **90**(6): p. 1107-1112.
84. Koh, S.S., et al., *Synergistic Enhancement of Nuclear Receptor Function by p160 Coactivators and Two Coactivators with Protein Methyltransferase Activities.* Journal of Biological Chemistry, 2001. **276**(2): p. 1089-1098.
85. Xu, J., et al., *Partial hormone resistance in mice with disruption of the steroid receptor coactivator-1 (SRC-1) gene.* Science, 1998. **279**(5358): p. 1922-5.
86. Han, S.J., et al., *Steroid receptor coactivator (SRC)-1 and SRC-3 differentially modulate tissue-specific activation functions of the progesterone receptor.* Mol Endocrinol, 2006. **20**(1): p. 45-55.
87. Shiozawa, T., et al., *Cyclic changes in the expression of steroid receptor coactivators and corepressors in the normal human endometrium.* J Clin Endocrinol Metab, 2003. **88**(2): p. 871-8.
88. Wieser, F., et al., *Endometrial nuclear receptor co-factors SRC-1 and N-CoR are increased in human endometrium during menstruation.* Mol Hum Reprod, 2002. **8**(7): p. 644-50.
89. Weiss, R.E., et al., *Mice deficient in the steroid receptor co-activator 1 (SRC-1) are resistant to thyroid hormone.* EMBO J, 1999. **18**(7): p. 1900-1904.
90. Yamada, T., et al., *SRC-1 is necessary for skeletal responses to sex hormones in both males and females.* J Bone Miner Res, 2004. **19**(9): p. 1452-61.
91. Khan, S.A., et al., *Estrogen Receptor Expression in Benign Breast Epithelium and Breast Cancer Risk.* Journal of the National Cancer Institute, 1998. **90**(1): p. 37-42.
92. Fullwood, M.J., et al., *An oestrogen-receptor-[agr]-bound human chromatin interactome.* Nature, 2009. **462**(7269): p. 58-64.
93. Marino, M. and P. Ascenzi, *Membrane association of estrogen receptor α and β influences 17β-estradiol-mediated cancer cell proliferation.* Steroids, 2008. **73**(9–10): p. 853-858.
94. Schiff, R., et al., *Cross-Talk between Estrogen Receptor and Growth Factor Pathways as a Molecular Target for Overcoming Endocrine Resistance.* Clinical Cancer Research, 2004. **10**(1): p. 331s-336s.
95. Kahlert, S., et al., *Estrogen Receptor α Rapidly Activates the IGF-1 Receptor Pathway.* Journal of Biological Chemistry, 2000. **275**(24): p. 18447-18453.
96. Wong, C.-W., et al., *Estrogen receptor-interacting protein that modulates its nongenomic activity-crosstalk with Src/Erk phosphorylation cascade.* Proceedings of the National Academy of Sciences, 2002. **99**(23): p. 14783-14788.
97. Levin, E. and R. Pietras, *Estrogen receptors outside the nucleus in breast cancer.* Breast Cancer Research and Treatment, 2008. **108**(3): p. 351-361.
98. Lannigan, D.A., *Estrogen receptor phosphorylation.* Steroids, 2003. **68**(1): p. 1-9.
99. Lemmon, M.A. and J. Schlessinger, *Cell Signaling by Receptor Tyrosine Kinases.* Cell, 2010. **141**(7): p. 1117-1134.
100. Wu, R.-C., C.L. Smith, and B.W. O'Malley, *Transcriptional Regulation by Steroid Receptor Coactivator Phosphorylation.* Endocrine Reviews, 2005. **26**(3): p. 393-399.

101. Lee, A.V., X. Cui, and S. Oesterreich, *Cross-Talk among Estrogen Receptor, Epidermal Growth Factor, and Insulin-like Growth Factor Signaling in Breast Cancer.* Clinical Cancer Research, 2001. **7**(12): p. 4429s-4435s.
102. Yarden, R.I., M.A. Wilson, and S.A. Chrysogelos, *Estrogen suppression of EGFR expression in breast cancer cells: A possible mechanism to modulate growth*.* Journal of Cellular Biochemistry, 2001. **81**(S36): p. 232-246.
103. Newman, S., Bates, NP., Vernimmen, D., Parker, MG., Hurst, HC., *Cofactor competition between the ligand-bound oestrogen receptor and an intron 1 enhancer leads to oestrogen repression of ERBB2 expression in breast cancer.* Oncogene, 2000. **19**(4): p. 490-497.
104. Lopez-Tarruella, S. and R. Schiff, *The Dynamics of Estrogen Receptor Status in Breast Cancer: Re-shaping the Paradigm.* Clinical Cancer Research, 2007. **13**(23): p. 6921-6925.
105. Cui, X., et al., *Insulin-Like Growth Factor-I Inhibits Progesterone Receptor Expression in Breast Cancer Cells via the Phosphatidylinositol 3-Kinase/Akt/Mammalian Target of Rapamycin Pathway: Progesterone Receptor as a Potential Indicator of Growth Factor Activity in Breast Cancer.* Molecular Endocrinology, 2003. **17**(4): p. 575-588.
106. Hudelist, G., et al., *Expression of Sex Steroid Receptors and their Co-Factors in Normal and Malignant Breast Tissue: AIB1 is a Carcinoma-Specific Co-Activator.* Breast Cancer Research and Treatment, 2003. **78**(2): p. 193-204.
107. Fleming, F.J., et al., *Differential Recruitment of Coregulator Proteins Steroid Receptor Coactivator-1 and Silencing Mediator for Retinoid and Thyroid Receptors to the Estrogen Receptor-Estrogen Response Element by β-Estradiol and 4-Hydroxytamoxifen in Human Breast Cancer.* Journal of Clinical Endocrinology & Metabolism, 2004. **89**(1): p. 375-383.
108. Wang, S., et al., *Disruption of the SRC-1 gene in mice suppresses breast cancer metastasis without affecting primary tumor formation.* Proceedings of the National Academy of Sciences, 2009. **106**(1): p. 151-156.
109. Leek, R.D., et al., *Association of Macrophage Infiltration with Angiogenesis and Prognosis in Invasive Breast Carcinoma.* Cancer Research, 1996. **56**(20): p. 4625-4629.
110. Wyckoff, J., et al., *A Paracrine Loop between Tumor Cells and Macrophages Is Required for Tumor Cell Migration in Mammary Tumors.* Cancer Research, 2004. **64**(19): p. 7022-7029.
111. Fleming, F.J., et al., *Expression of SRC-1, AIB1, and PEA3 in HER2 mediated endocrine resistant breast cancer; a predictive role for SRC-1.* Journal of Clinical Pathology, 2004. **57**(10): p. 1069-1074.
112. Al-azawi, D., et al., *Ets-2 and p160 proteins collaborate to regulate c-Myc in endocrine resistant breast cancer.* Oncogene, 2008. **27**(21): p. 3021-3031.
113. McBryan, J., et al., *Metastatic progression with resistance to aromatase inhibitors is driven by the steroid receptor coactivator SRC-1.* Cancer Research, 2011.
114. Han, J. and D. Crowe, *Steroid receptor coactivator 1 deficiency increases MMTV-neu mediated tumor latency and differentiation specific gene expression, decreases metastasis, and inhibits response to PPAR ligands.* BMC Cancer, 2010. **10**(1): p. 629.
115. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition.* The Journal of Clinical Investigation, 2009. **119**(6): p. 1420-1428.
116. Yang, J., et al., *Twist, a Master Regulator of Morphogenesis, Plays an Essential Role in Tumor Metastasis.* Cell, 2004. **117**(7): p. 927-939.
117. Hong, J., et al., *Phosphorylation of Serine 68 of Twist1 by MAPKs Stabilizes Twist1 Protein and Promotes Breast Cancer Cell Invasiveness.* Cancer Research, 2011. **71**(11): p. 3980-3990.

118. Luo, B.-H., C.V. Carman, and T.A. Springer, *Structural Basis of Integrin Regulation and Signaling.* Annual Review of Immunology, 2007. **25**(1): p. 619-647.
119. Desgrosellier, J.S. and D.A. Cheresh, *Integrins in cancer: biological implications and therapeutic opportunities.* Nat Rev Cancer, 2010. **10**(1): p. 9-22.
120. Wakeling, A.E., M. Dukes, and J. Bowler, *A Potent Specific Pure Antiestrogen with Clinical Potential.* Cancer Research, 1991. **51**(15): p. 3867-3873.
121. Bowler, J., et al., *Novel steroidal pure antiestrogens.* Steroids, 1989. **54**(1): p. 71-99.
122. Osborne, C.K., et al., *Comparison of the Effects of a Pure Steroidal antiestrogen With Those of Tamoxifen in a Model of Human Breast Cancer.* Journal of the National Cancer Institute, 1995. **87**(10): p. 746-750.
123. Howell, A., et al., *Response to a specific antioestrogen (ICI 182780) in tamoxifen-resistant breast cancer.* The Lancet, 1995. **345**(8941): p. 29-30.
124. Young, O.E., et al., *Effects of fulvestrant 750mg in premenopausal women with oestrogen-receptor-positive primary breast cancer.* European journal of cancer (Oxford, England : 1990), 2008. **44**(3): p. 391-399.
125. Gutteridge, E., Robertson, JFR., Cheung, KL. , *Effects of fulvestrant on estrogen receptor levels during long-term treatment of patients with advanced breast cancer - final results.* . Breast Cancer Res Treat. , 2004. **88** p. suppl 1:S177.
126. Robertson, J.F.R., et al., *Activity of Fulvestrant 500 mg Versus Anastrozole 1 mg As First-Line Treatment for Advanced Breast Cancer: Results From the FIRST Study.* Journal of Clinical Oncology, 2009. **27**(27): p. 4530-4535.
127. Di Leo, A., et al., *Results of the CONFIRM Phase III Trial Comparing Fulvestrant 250 mg With Fulvestrant 500 mg in Postmenopausal Women With Estrogen Receptor–Positive Advanced Breast Cancer.* Journal of Clinical Oncology, 2010. **28**(30): p. 4594-4600.
128. McDonnell, D.P. and S.E. Wardell, *The molecular mechanisms underlying the pharmacological actions of ER modulators: implications for new drug discovery in breast cancer.* Current Opinion in Pharmacology, 2010. **10**(6): p. 620-628.
129. Evans, C.T., et al., *Isolation and characterization of a complementary DNA specific for human aromatase-system cytochrome P-450 mRNA.* Proceedings of the National Academy of Sciences, 1986. **83**(17): p. 6387-6391.
130. Miller, W.R., R.A. Hawkins, and A.P.M. Forrest, *Significance of Aromatase Activity in Human Breast Cancer.* Cancer Research, 1982. **42**(8 Supplement): p. 3365s-3368s.
131. Smith, I.E. and M. Dowsett, *Aromatase Inhibitors in Breast Cancer.* New England Journal of Medicine, 2003. **348**(24): p. 2431-2442.
132. Thijssen, J.H.H. and M.A. Blankenstein, *Endogenous oestrogens and androgens in normal and malignant endometrial and mammary tissues.* European Journal of Cancer and Clinical Oncology, 1989. **25**(12): p. 1953-1959.
133. Stein, R., Dowsett, M., Hedley, A., Gazet, JC., Ford, HT., Coombes, RC. , *The clinical and endocrine effects of 4-hydroxyandrostenedione alone and in combination with oserelin in premenopausal women with advanced breast cancer.* Br J Cancer, 1990. **62**(4): p. 679-683.
134. Harper, M.J.K. and A.L. Walpole, *Contrasting Endocrine Activities of cis and trans Isomers in a Series of Substituted Triphenylethylenes.* Nature, 1966. **212**(5057): p. 87-87.
135. Jordan, V.C., *Tamoxifen: a most unlikely pioneering medicine.* Nat Rev Drug Discov, 2003. **2**(3): p. 205-213.
136. EMMENS, C.W., *Postcoital Contraception.* British Medical Bulletin, 1970. **26**(1): p. 45-51.

137. Ali, S., L. Buluwela, and R.C. Coombes, *Antiestrogens and Their Therapeutic Applications in Breast Cancer and Other Diseases.* Annual Review of Medicine, 2011. **62**(1): p. 217-232.
138. Desta, Z., et al., *Comprehensive Evaluation of Tamoxifen Sequential Biotransformation by the Human Cytochrome P450 System in Vitro: Prominent Roles for CYP3A and CYP2D6.* Journal of Pharmacology and Experimental Therapeutics, 2004. **310**(3): p. 1062-1075.
139. McDonnell, D.P., et al., *Analysis of estrogen receptor function in vitro reveals three distinct classes of antiestrogens.* Molecular Endocrinology, 1995. **9**(6): p. 659-69.
140. Berry, M., D. Metzger, and P. Chambon, *Role of the two activating domains of the oestrogen receptor in the cell-type and promoter-context dependent agonistic activity of the anti-oestrogen 4-hydroxytamoxifen.* EMBO J., 1990. **9**: p. 2811-2818.
141. Thomas, R.S., et al., *Phosphorylation at serines 104 and 106 by Erk1/2 MAPK is important for estrogen receptor-α activity.* Journal of Molecular Endocrinology, 2008. **40**(4): p. 173-184.
142. McDonnell, D.P., *The Molecular Pharmacology of SERMs.* Trends in Endocrinology & Metabolism, 1999. **10**(8): p. 301-311.
143. Smith, C.L., Z. Nawaz, and B.W. O'Malley, *Coactivator and corepressor regulation of the agonist/antagonist activity of the mixed antiestrogen,4-hydroxytamoxifen.* Mol Endocrinol, 1997. **11**: p. 657-666.
144. Shang, Y. and M. Brown, *Molecular Determinants for the tissue specificty of SERMs* Science, 2002. **295**: p. 2465-2468.
145. Webb, P., et al., *Estrogen Receptor Activation Function 1 Works by Binding p160 Coactivator Proteins.* Molecular Endocrinology, 1998. **12**(10): p. 1605-1618.
146. Kressler, D., M.B. Hock, and A. Kralli, *Coactivators PGC-1β and SRC-1 Interact Functionally to Promote the Agonist Activity of the Selective Estrogen Receptor Modulator Tamoxifen.* Journal of Biological Chemistry, 2007. **282**(37): p. 26897-26907.
147. McDonnell, D.P. and J.D. Norris, *Analysis of the molecular pharmacology of estrogen receptor agonists and antagonists provides insights into the mechanism of action of estrogen in bone.* Osteoporosis International, 1997. **7**(0): p. 29-34.
148. Chia, S.K. and A.C. Wolff, *With maturity comes confidence: EBCTCG tamoxifen update.* The Lancet. **378**(9793): p. 747-749.
149. EBCTCG, *Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials.* The Lancet, 2005. **365**(9472): p. 1687-1717.
150. (EBCTCG), E.B.C.T.C.G., *Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials,.* Lancet, 2005. **365**: p. 1687–1717.
151. Ottaviano, Y.L., et al., *Methylation of the Estrogen Receptor Gene CpG Island Marks Loss of Estrogen Receptor Expression in Human Breast Cancer Cells.* Cancer Research, 1994. **54**(10): p. 2552-2555.
152. Yang, X., et al., *Synergistic Activation of Functional Estrogen Receptor (ER)-α by DNA Methyltransferase and Histone Deacetylase Inhibition in Human ER-α-negative Breast Cancer Cells.* Cancer Research, 2001. **61**(19): p. 7025-7029.
153. Fan, J., et al., *ERα negative breast cancer cells restore response to endocrine therapy by combination treatment with both HDAC inhibitor and DNMT inhibitor.* Journal of Cancer Research and Clinical Oncology, 2008. **134**(8): p. 883-890.
154. Branković-Magić, M.B.-M., et al., *Progesterone receptor status of breast cancer metastases.* Journal of Cancer Research and Clinical Oncology, 2002. **128**(1): p. 55-60.

155. Gross, G.E., et al., *Multiple Progesterone Receptor Assays in Human Breast Cancer.* Cancer Research, 1984. **44**(2): p. 836-840.
156. Cui, X., et al., *Biology of Progesterone Receptor Loss in Breast Cancer and Its Implications for Endocrine Therapy.* Journal of Clinical Oncology, 2005. **23**(30): p. 7721-7735.
157. Normanno, N., et al., *Epidermal growth factor-related peptides in the pathogenesis of human breast cancer.* Breast Cancer Research and Treatment, 1994. **29**(1): p. 11-27.
158. Encarnación, C.A., et al., *Measurement of steroid hormone receptors in breast cancer patients on tamoxifen.* Breast Cancer Research and Treatment, 1993. **26**(3): p. 237-246.
159. Gee, J.M.W., et al., *The Antiepidermal Growth Factor Receptor Agent Gefitinib (ZD1839/Iressa) Improves Antihormone Response and Prevents Development of Resistance in Breast Cancer in Vitro.* Endocrinology, 2003. **144**(11): p. 5105-5117.
160. Knowlden, J.M., et al., *Elevated Levels of Epidermal Growth Factor Receptor/c-erbB2 Heterodimers Mediate an Autocrine Growth Regulatory Pathway in Tamoxifen-Resistant MCF-7 Cells.* Endocrinology, 2003. **144**(3): p. 1032-1044.
161. Shou, J., et al., *Mechanisms of Tamoxifen Resistance: Increased Estrogen Receptor-HER2/neu Cross-Talk in ER/HER2–Positive Breast Cancer.* Journal of the National Cancer Institute, 2004. **96**(12): p. 926-935.
162. Osborne, C.K., et al., *Role of the Estrogen Receptor Coactivator AIB1 (SRC-3) and HER-2/neu in Tamoxifen Resistance in Breast Cancer.* Journal of the National Cancer Institute, 2003. **95**(5): p. 353-361.
163. Lavinsky, R.M., et al., *Diverse signaling pathways modulate nuclear receptor recruitment of N-CoR and SMRT complexes.* Proceedings of the National Academy of Sciences, 1998. **95**(6): p. 2920-2925.
164. Romano, A., et al., *Identification of novel ER-α target genes in breast cancer cells: Gene- and cell-selective co-regulator recruitment at target promoters determines the response to 17β-estradiol and tamoxifen.* Molecular and Cellular Endocrinology, 2010. **314**(1): p. 90-100.
165. Liu, Z., et al., *Coactivator/corepressor ratios modulate PR-mediated transcription by the selective receptor modulator RU486.* Proceedings of the National Academy of Sciences, 2002. **99**(12): p. 7940-7944.
166. Lonard, D.M., S.Y. Tsai, and B.W. O'Malley, *Selective Estrogen Receptor Modulators 4-Hydroxytamoxifen and Raloxifene Impact the Stability and Function of SRC-1 and SRC-3 Coactivator Proteins.* Molecular and Cellular Biology, 2004. **24**(1): p. 14-24.
167. Haugan Moi, L.L., et al., *Effect of Low-Dose Tamoxifen on Steroid Receptor Coactivator 3/Amplified in Breast Cancer 1 in Normal and Malignant Human Breast Tissue.* Clinical Cancer Research, 2010. **16**(7): p. 2176-2186.
168. Scott, D., et al., *Changes in expression of steroid receptors, their downstream target genes and their associated co-regulators during the sequential acquisition of tamoxifen resistance in vitro.* International Journal of Oncology., 2007. **31**(3): p. 557-565.
169. Redmond, A.M., et al., *Coassociation of Estrogen Receptor and p160 Proteins Predicts Resistance to Endocrine Treatment; SRC-1 is an Independent Predictor of Breast Cancer Recurrence.* Clinical Cancer Research, 2009. **15**(6): p. 2098-2106.
170. Zwart, W., et al., *PKA-induced resistance to tamoxifen is associated with an altered orientation of ER[alpha] towards co-activator SRC-1.* EMBO J, 2007. **26**(15): p. 3534-3544.

171. Holm, C., et al., *Association Between Pak1 Expression and Subcellular Localization and Tamoxifen Resistance in Breast Cancer Patients.* Journal of the National Cancer Institute, 2006. **98**(10): p. 671-680.
172. McCartan, D., et al., *Global characterization of the SRC-1 transcriptome identifies ADAM22 as an ER-independent mediator of endocrine resistant breast cancer.* Cancer Research, 2011.
173. Lewis, M.T., *Homeobox genes in mammary gland development and neoplasia.* Breast Cancer Res, 2000. **2**: p. 158-169.
174. McCoy, E., et al., *Six1 expands the mouse mammary epithelial stem/progenitor cell pool and induces mammary tumours that undergo epithelial mesenchymal transition.* . The Journal of Clinical Investigation 2009. **119**: p. 2663-2667.
175. Chen, H. and S. Sukumar, *Role of homeobox genes in normal mammary gland development and breast tumourigenesis.* Journal of Mammary Gland Biology and Neoplasia, 2003. **8**: p. 159-175.
176. Visvader, J. and G. Lindeman, *Transcriptional regulators in mammary gland development and cancer.* . The International Journal of Biochemistry & Cell Biology 2003. **35**: p. 1034-1051.
177. Shah, N. and S. Sukumar, *The Hox genes and their role in oncogenesis.* . Nature Reviews Cancer, 2010. **10**: p. 361-370.
178. Chen, F. and M.R. Capecchi, *Paralogous mouse Hox genes, Hoxa9, Hoxb9, and Hoxd9, function together to control development of the mammary gland in response to pregnancy.* PNAS, 1999. **96**: p. 541-546.
179. Garcia-Gasca, A. and D.D. Spyropoulos, *Differential mammary morphogenesis along the anteroposterior axis in Hoxc6 gene targeted mice.* Development Dynamics, 2000. **219**: p. 261-276.
180. Pearson, J.C., D. Lemons, and W. McGinnis, *Modulating Hox gene functions during animal body patterning.* National Review Genetics., 2005. **6**: p. 893-904.
181. Jina, K., et al., *The HOXB7 protein renders breast cancer cells resistant to tamoxifen through activation of the EGFR pathway.* PNAS, 2011.
182. Ma, X., et al., *A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen.* . Cancer Cell, 2004. **5**: p. 607-616.
183. Jervall, P.L., et al., *Exploring the two-gene ratio in breast cancer - independent roles for HOXB13 and IL17BR in prediction of clinical outcome.* Breast Cancer Res Treat., 2008. **107**(225-234).
184. Raman, V., et al., *Compromised HOXA5 function can limit p53 expression in human breast tumours.* Nature, 2000. **405**: p. 974-978.
185. Chu, M.C., F.B. Selam, and H.S. Taylor, *HOXA10 regulates p53 expression and matriel invasion in human breast cancer cells.* . Cancer Biology Therapeutics, 2004. **3**: p. 568-572.
186. Zhang, X., et al., *HOXC6 and HOXC11 increase transcription of S100beta gene in BrdU-induced in vitro differentiation of GOTO neuroblastoma cells into Schwannian cells.* J Cell Mol Med., 2007. **11**(2): p. 299-306.
187. BRONZERT, D.A., G.L. GREENE, and M.E. LIPPMAN, *Selection and Characterization of a Breast Cancer Cell Line Resistant to the Antiestrogen LY 117018.* Endocrinology, 1985. **117**(4): p. 1409-1417.
188. Coezy, E., J.-L. Borgna, and H. Rochefort, *Tamoxifen and Metabolites in MCF7 Cells: Correlation between Binding to Estrogen Receptor and Inhibition of Cell Growth.* Cancer Research, 1982. **42**(1): p. 317-323.
189. Fire, A., et al., *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.* Nature, 1998. **391**(6669): p. 806-811.

190. Chesnoy, S. and L. Huang, *Structure and function of lipid-DNA complexes for gene delivery.* Annual Review of Biophysics and Biomolecular Structure, 2000. **29**(1): p. 27-47.
191. Fraga, D., T. Meulia, and S. Fenster, *Real-Time PCR*, in *Current Protocols Essential Laboratory Techniques*. 2008, John Wiley & Sons, Inc.
192. Livak, K.J. and T.D. Schmittgen, *Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method.* Methods, 2001. **25**(4): p. 402-408.
193. Allred, D.C., et al., *Association of p53 Protein Expression With Tumor Cell Proliferation Rate and Clinical Outcome in Node-Negative Breast Cancer.* Journal of the National Cancer Institute, 1993. **85**(3): p. 200-206.
194. Das, P.M. and R. Singal, *DNA Methylation and Cancer.* Journal of Clinical Oncology, 2004. **22**(22): p. 4632-4642.
195. Tazi, J. and A. Bird, *Alternative chromatin structure at CpG islands.* Cell, 1990. **60**(6): p. 909-920.
196. Weber, M., et al., *Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells.* Nat Genet, 2005. **37**(8): p. 853-862.
197. Langmead, B., Trapnell, C., Pop, M., Salzberg, SL. , *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* . Genome Biol. , 2009. **10**(3): p. R25
198. Zhang, Y., Liu, T., Meyer, CA., Eeckhoute, J., Johnson, DS., Bernstein, BE., Nusbaum, C., Myers, RM., Brown, M., Li, W., Liu, XS. , *Model-based analysis of ChiP-Seq (MACS).* . Genome Biol, 2008. **9**: p. R137
199. Zwart, W., et al., *Oestrogen receptor-co-factor-chromatin specificity in the transcriptional regulation of breast cancer.* EMBO J, 2011. **30**(23): p. 4764-4776.
200. Ross-Innes, C., Stark, R., Holmes, KA., Schmidt, D., Spyrou, C., Russell, R., Massie, CE., Vowler, SL., Eldridge, M., Carroll, JS. , *Cooperative interaction between retinoic acid receptor-alpha and estrogen receptor in breast cancer.* . Genes & Development, 2010. **24**: p. 171-182
201. Zhu, L., Pages, H., Gazin, C., Lawson, N., Lin, S., Lapointe, D., Green, M. , *ChIPpeakAnno: Batch annotation of the peaks identified from either ChIP-seq or ChIP-chip experiments.* . 2011 **R package version 1.8.0.** .
202. Salmon-Divon, M., Dvinge, H., Tammoja, K., Bertone, P. , *PeakAnalyzer: Genome-wide annotation of chromatin binding and modification loci.* . BMC Bioinformatics 2010.
203. Shin, H., Liu, T., Manrai, AK., Liu, XS. , *CEAS: cis-regulatory element annotation system.* . Bioionformatics, 2009 **25**(19): p. 2605-2606
204. Robasky, K., Bulyk, ML. , *UniPROBE: an online database of protein binding microarray data on protein-DNA interactions.* . Nucleic Acids Res. , 2009 **37(database issue):** : p. D77-82
205. Schwartz, S., Elnitski, L., Li, M., Weirauch, M., Riemer, C., Smit, A., Green, ED., Hardison, RC., Miller, W. , *MultiPipMaker and supporting tools: alignments and analysis of multiple genomic DNA sequences.* Nucleic Acids Res 2003. **31**: p. 3518-3524.
206. Quinlan, A.a.H., IM, , *BEDTools: a flexible suite of utilities for comparing genomic features.* . Bioinformatics, 2010. **26** (6): p. 841–842.
207. Kristiansen, G., M. Sammar, and P. Altevogt, *Tumour biological aspects of CD24, a mucin-like adhesion molecule.* . J Mol Histol, 2004. **35**: p. 255-262.
208. Zarn, J., Jackson,DG., Bell, MV., Jones, T., Weber, E., Sheer, D., Waibel, R., Stahel, RA., *The small cell lung cancer antigen cluster-4 and the leukocyte antigen CD24 are*

allelic isoforms of the same gene (CD24) on chromosome band 6q21. Cytogenet Cell Genet, 1995. **70**(1-2): p. 119-125.

209. Nielsen, P.J., et al., *Altered Erythrocytes and a Leaky Block in B-Cell Development in CD24/HSA-Deficient Mice.* Blood, 1997. **89**(3): p. 1058-1067.
210. Aigner, S., Ruppert, M., Hubbe, M., Sammar, M., Sthoeger, Z., Butcher, EC., Vestweber, D., Altevogt, P., *Heat stable antigen (mouse CD24) supports myeloid cell binding to endothelial and platelet P-selectin.* Int Immunol, 1995. **7**: p. 1557-1565.
211. Jacobson, K. and C. Dietrich, *Looking at lipid rafts?* Trends in cell biology, 1999. **9**(3): p. 87-91.
212. Zarn, J.A., et al., *Association of CD24 with the Kinase c-fgr in a Small Cell Lung Cancer Cell Line and with the Kinase lyn in an Erythroleukemia Cell Line.* Biochemical and Biophysical Research Communications, 1996. **225**(2): p. 384-391.
213. Aigner, S., et al., *CD24 mediates rolling of breast carcinoma cells on P-selectin.* The FASEB Journal, 1998. **12**(12): p. 1241-1251.
214. Kim, Y.J., et al., *P-selectin deficiency attenuates tumor growth and metastasis.* Proceedings of the National Academy of Sciences, 1998. **95**(16): p. 9325-9330.
215. Kristiansen, G., et al., *CD24 Expression Is a New Prognostic Marker in Breast Cancer.* Clinical Cancer Research, 2003. **9**(13): p. 4906-4913.
216. Schindelmann S, W.J., Grundmann R, Kreienberg R, Zeillinger R, Deissler H, *Expression profiling of mammary carcinoma cell lines: Correlation of in vitro invasiveness with expression of CD24.* Tumour Biol 2002. **23**: p. 139–145.
217. Al-Hajj, M., et al., *Prospective identification of tumorigenic breast cancer cells.* Proc Natl Acad Sci USA, 2003. **100**: p. 3983 - 3988.
218. Al-Hajj, M., et al., *Prospective identification of tumorigenic breast cancer cells.* Proceedings of the National Academy of Sciences, 2003. **100**(7): p. 3983-3988.
219. Blick, T., et al., *Epithelial Mesenchymal Transition Traits in Human Breast Cancer Cell Lines Parallel the $CD44^{hi/}CD24^{lo/-}$ Stem Cell Phenotype in Human Breast Cancer.* Journal of Mammary Gland Biology and Neoplasia, 2010. **15**(2): p. 235-252.
220. Kaipparettu, B.A., et al., *Estrogen-mediated downregulation of CD24 in breast cancer cells.* International Journal of Cancer, 2008. **123**(1): p. 66-72.
221. Vesuna, F., et al., *Twist Modulates Breast Cancer Stem Cells by Transcriptional Regulation of CD24 Expression.* Neoplasia, 2009. **11**(12): p. 1318-1328.
222. Sørlie, T., *Molecular portraits of breast cancer: tumour subtypes as distinct disease entities.* European journal of cancer (Oxford, England : 1990), 2004. **40**(18): p. 2667-2675.
223. Christensen, K.L., et al., *Chapter 5 The Six Family of Homeobox Genes in Development and Cancer*, in *Advances in Cancer Research*, F.V.W. George and K. George, Editors. 2008, Academic Press. p. 93-126.
224. Yu, C., et al., *An Essential Function of the SRC-3 Coactivator in Suppression of Cytokine mRNA Translation and Inflammatory Response.* Molecular cell, 2007. **25**(5): p. 765-778.
225. Johnstone, R., See, RH., Sells, SF., Wang, J., Muthukkumar, S., Englert, C., Haber, DA., Licht, JD., Sugrue, SP., Roberts, T., Rangnekar, VM., Shi, Y. , *A novel repressor, par-4, modulates transcription and growth suppression functions of the Wilms' tumor suppressor WT1.* . Mol Cell Biol, 1996. **16**: p. 6945–56.
226. Díaz-Meco, M.T., et al., *The Product of par-4, a Gene Induced during Apoptosis, Interacts Selectively with the Atypical Isoforms of Protein Kinase C.* Cell, 1996. **86**(5): p. 777-786.
227. Johnstone, R.W., et al., *Mapping of the Human PAWR (par-4) Gene to Chromosome 12q21.* Genomics, 1998. **53**(2): p. 241-243.

228. Ranganathan, P. and V.M. Rangnekar, *Regulation of Cancer Cell Survival by Par-4.* Annals of the New York Academy of Sciences, 2005. **1059**(1): p. 76-85.
229. Cheema, S.K., et al., *Par-4 Transcriptionally Regulates Bcl-2 through a WT1-binding Site on the bcl-2 Promoter.* Journal of Biological Chemistry, 2003. **278**(22): p. 19995-20005.
230. Garcia-Cao, I., et al., *Genetic inactivation of Par4 results in hyperactivation of NF-[kappa]B and impairment of JNK and p38.* EMBO Rep, 2003. **4**(3): p. 307-312.
231. Page, G., Kögel, D., Rangnekar, V., Scheidtmann, KH: , *Interaction partners of Dlk/ZIP kinase: co-expression of Dlk/ZIP kinase and Par-4 results in cytoplasmic retention and apoptosis.* Oncogene, 1999. **18**: p. 7265-7273.
232. Qiu, G., Ahmed, M., Sells, SF., Mohiuddin, M., Weinstein, MH., Rangnekar, VM: , *Mutually exclusive expression patterns of Bcl-2 and Par-4 in human prostate tumors consistent with down-regulation of Bcl-2 by Par-4.* . Oncogene 1999. **18**: p. 623-631.
233. Chakraborty, M., et al., *Par-4 Drives Trafficking and Activation of Fas and FasL to Induce Prostate Cancer Cell Apoptosis and Tumor Regression.* Cancer Research, 2001. **61**(19): p. 7255-7263.
234. Boghaert, E., et al., *Immunohistochemical analysis of the proapoptotic protein Par-4 in normal rat tissues.* Cell Growth Differ, 1997. **8**(8): p. 881-890.
235. Bieberich, E., et al., *Regulation of cell death in mitotic neural progenitor cells by asymmetric distribution of prostate apoptosis response 4 (PAR-4) and simultaneous elevation of endogenous ceramide.* The Journal of Cell Biology, 2003. **162**(3): p. 469-479.
236. Cook, J., Krishnan, S., Ananth, S., Sells, SF., Shi, Y., Walther, MM., Linehan, WM., Sukhatme, VP., Weinstein, MH., Rangnekar, VM. , *Decreased expression of the pro-apoptotic protein Par-4 in renal cell carcinoma.* . Oncogene, 1999. **18**: p. 1205–8.
237. Kogel, D., et al., *Dlk/ZIP kinase-induced apoptosis in human medulloblastoma cells: requirement of the mitochondrial apoptosis pathway.* Br J Cancer, 2001. **85**(11): p. 1801-1808.
238. Boehrer, S., Chow, KU., Puccetti, E., Ruthardt, M., Godzisard, S., Krapohl, A., Schneider, B., Hoelzer, D., Mitrou, PS., Rangnekar, VM., Weidmann, E. , *Deregulated expression of prostate apoptosis response gene-4 in less differentiated lymphocytes and inverse expressional patterns of par-4 and bcl-2 in acute lymphocytic leukemia* Hematol J, 2001. **2**(103–7).
239. Moreno-Bueno, G., et al., *Inactivation of the Candidate Tumor Suppressor Par-4 in Endometrial Cancer.* Cancer Research, 2007. **67**(5): p. 1927-1934.
240. Garcia-Cao, I., et al., *Tumour-suppression activity of the proapoptotic regulator Par4.* EMBO Rep, 2005. **6**(6): p. 577-583.
241. El-Guendy, N., et al., *Identification of a Unique Core Domain of Par-4 Sufficient for Selective Apoptosis Induction in Cancer Cells.* Molecular and Cellular Biology, 2003. **23**(16): p. 5516-5525.
242. Shareef, M.M., et al., *Role of Tumor Necrosis Factor-α and TRAIL in High-Dose Radiation–Induced Bystander Signaling in Lung Adenocarcinoma.* Cancer Research, 2007. **67**(24): p. 11811-11820.
243. Kimura, K. and E.P. Gelmann, *Tumor Necrosis Factor-α and Fas Activate Complementary Fas-associated Death Domain-dependent Pathways That Enhance Apoptosis Induced by γ-Irradiation.* Journal of Biological Chemistry, 2000. **275**(12): p. 8610-8617.
244. Goswami, A., P. Ranganathan, and V.M. Rangnekar, *The Phosphoinositide 3-Kinase/Akt1/Par-4 Axis: A Cancer-Selective Therapeutic Target.* Cancer Research, 2006. **66**(6): p. 2889-2892.

245. Goswami, A., et al., *Binding and Phosphorylation of Par-4 by Akt Is Essential for Cancer Cell Survival.* Molecular cell, 2005. **20**(1): p. 33-44.
246. Nagai, M., Gerhard, R., Salaorni, S., Fregnani, JH., Nonogaki, S., Netto, MM., Soares, FA., *Down-regulation of the candidate tumor suppressor gene PAR-4 is associated with poor prognosis in breast cancer.* Int J Oncol, 2010. **37**(1): p. 41-9.
247. Casolari, D., Pereira, MC., deBessa Garcia, SA., Nagai, MA., *Insulin-like growth factor-1 and 17β-estradiol down-regulate prostate apoptosis response-4 expression in MCF-7 breast cancer cells.* Int J Mol Med, 2011. **28**(3): p. 337-342.
248. Acharya, M., et al., *Human PRKC Apoptosis WT1 Regulator Is a Novel PITX2-interacting Protein That Regulates PITX2 Transcriptional Activity in Ocular Cells.* Journal of Biological Chemistry, 2009. **284**(50): p. 34829-34838.
249. O'Brien, C., et al., *Breast Cancer Stem Cells and Their Role in Resistance to Endocrine Therapy.* Hormones and Cancer, 2011. **2**(2): p. 91-103.
250. Singh, A. and J. Settleman, *EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer.* Oncogene, 2010. **29**(34): p. 4741-4751.
251. Bieberich, E., et al., *Selective apoptosis of pluripotent mouse and human stem cells by novel ceramide analogues prevents teratoma formation and enriches for neural precursors in ES cell–derived neural transplants.* The Journal of Cell Biology, 2004. **167**(4): p. 723-734.
252. Miranda, T.B. and P.A. Jones, *DNA methylation: The nuts and bolts of repression.* Journal of Cellular Physiology, 2007. **213**(2): p. 384-390.
253. Chen, Z.-x. and A.D. Riggs, *DNA Methylation and Demethylation in Mammals.* Journal of Biological Chemistry, 2011. **286**(21): p. 18347-18353.
254. Baylin, S.B., *DNA methylation and gene silencing in cancer.* Nature Clinical Practice Oncology, 2005. **2**(S4-S11).
255. Yoon, H.-G., et al., *N-CoR Mediates DNA Methylation-Dependent Repression through a Methyl CpG Binding Protein Kaiso.* Molecular cell, 2003. **12**(3): p. 723-734.
256. Riau, A.K., et al., *Aberrant DNA Methylation of Matrix Remodeling and Cell Adhesion Related Genes in Pterygium.* PLoS ONE, 2011. **6**(2): p. e14687.
257. Rojo, F., *Mechanisms of transcriptional repression.* Current Opinion in Microbiology, 2001. **4**(2): p. 145-151.
258. Prat, A., M.J. Ellis, and C.M. Perou, *Practical implications of gene-expression-based assays for breast oncologists.* Nat Rev Clin Oncol, 2012. **9**(1): p. 48-57.

Appendix

Table 7.1. Buffers used in Western Blotting

RIPA Lysis Buffer	10X Running Buffer	10X Wet Transfer Buffer
150mM NaCl	1.92M Glycine	192mM Glycine
1% NP-40	250mM Trizma Base	250mM Trizma Base
0.5% Na deoxycholate	1% SDS	
0.1% SDS		
50 mM Tris, pH 8.0		

Table 7.2. Primary Antibodies used in FACS analysis

Antigen	Antibody	Catalogue number
CD24	PE mouse anti human	BD555428
CD44	FITC mouse anti human	BD555478
CD49f	FITC rat anti human	BD555735
EpCAM	APC mouse anti human	347200
IgG2a κ Isotype Control	APC Mouse	551414
IgG2a κ Isotype Control	PE Mouse	BD555574
IgG2b κ Isotype Control	FITC Mouse	BD555742
IgG2a κ Isotype Control	FITC Rat	BD555843
Anti Mouse Ig, κNegative Control (FBS) Compensation Particles		BD552843
Anti Rat Ig, κNegative Control (FBS) Compensation Particles		BD552844

All antibodies and controls are supplied by BD Biosciences

Table 7.3. Buffers used in Chromatin Immunoprecipitation

Lysis Buffer 1	Lysis Buffer 2	Lysis Buffer 3
50mM Hepes-KOH, pH 7.5	10mM Tris-HCL, pH 8	10mM Tris-HCL, pH 8
140mM NaCl	200mM NaCl	100m NaCl
1mM EDTA	1mM EDTA	1mM EDTA
10% Glycerol	1mM EGTA	1mM EGTA
0.5% NP-40/gepal CA630	Double distilled water	0.1% Na-Deoxycholate
0.25% Triton X-100		0.5% N-lauroylsarcosine
Double distilled water		Double distilled water

RIPA Wash Buffer	TE Buffer	Elution buffer
50mM Hepes, pH 8	10mM Tris-Cl, pH 7.5	1M Tris, pH 8
0.5M EDTA	1mM EDTA	0.5M EDTA
10% NP-40		10%SDS
10% Na-deoxycholate		Double distilled water
Double distilled water		
8M Lithium Chloride		
½ protease inhibitor tablet		

Proteinase K mix
1X TE buffer
10mg/mL Glycogen
20mg/mL Proteinase K

Table 7.4. List of 1,061 SRC-1 downregulated genes

Gene Name	Chromosome	Fold Enrichment
TLE1	chr9	300.53
ZNF746	chr7	294.29
HIRA	chr22	261.11
SNX19	chr11	245.86
PGS1	chr17	216.35
DDX28	chr16	198.58
NDUFA11	chr19	174.11
CHD9	chr16	155.17
FSTL5	chr4	155.17
INSM2	chr14	142.24
SLC22A23	chr6	140.49
SMG6	chr17	140.49
HIRA	chr22	137.39
RFX1	chr19	132.87
NCAPH2	chr22	131.71
NFKBIZ	chr3	129.31
C9orf9	chr9	126.08
MAF1	chr8	125.91
FAM98B	chr15	125.03
SPRYD4	chr12	123.92
CD24	**chrY**	**122.93**
FAM162A	chr3	122.93
POU2F1	chr1	122.93
SMEK2	chr2	122.85
KLC4	chr6	121.23
MFN2	chr1	116.38
FRAT2	chr10	114.15
MAPK8IP2	chr22	114.15
U2AF1	chr21	113.43
HIC1	chr17	113.15
NAPEPLD	chr7	112.5
STX16	chr20	110.31
TSJD2	chr6	108.74
CCDC130	chr19	105.37
DGKQ	chr4	105.37
TTYH2	chr17	105.3
TRAPPC10	chr21	105.37
B2M	chr15	105.07
SMPD1	chr11	105.07
NADK	chr1	104.44
ATXN2L	chr16	104.06
GPBP1	chr5	103.45
RPL36	chr19	103.45
MTHFSD	chr16	101.56
CELSR3	chr3	101.2
NFKBIZ	chr3	101.2
MAF1	chr8	100.11
STRN	chr2	97.06
AGRN	chr1	96.98
CCRN4L	chr4	96.98
CNOT7	chr8	96.98
DNAJC30	chr7	96.98
HEXDC	chr17	96.98
KIAA0430	chr16	96.98
NDE1	chr16	96.98
NUP93	chr16	96.98
PPP1R3B	chr8	96.98

Gene Name	Chromosome	Fold Enrichment
RPS15	chr19	96.98
ZNF165	chr6	96.98
CHP	chr15	96.59
KLF3	chr4	96.59
RUNX1	chr21	96.59
USP4	chr3	96.59
POLR2A	chr17	94.21
C17orf44	chr17	94.04
NECAP2	chr1	92
C20orf24	chr20	90.06
PSMA1	chr11	89.61
ICA1	chr7	89.22
FAM149B1	chr10	88.7
SLC35E1	chr19	88.67
ANXA2	chr15	88.25
PSMA3	chr14	88.17
C22orf32	chr22	87.81
CCDC7	chr10	87.81
HIST3H2A	chr1	87.81
HSPA6	chr1	87.81
IRS2	chr13	87.81
PPAPDC2	chr9	87.81
RUFY3	chr4	87.81
TMEM63A	chr1	87.81
DMAP1	chr1	87.75
FXR2	chr17	87.28
HEXDC	chr17	87.28
TRIB1	chr8	86.9
MALAT1	chr11	86.21
HEXA	chr15	86.12
TIA1	chr2	85.9
C17orf44	chr17	85.57
TMEM39A	chr3	85.35
CTDSP2	chr12	85.23
USP42	chr7	84.86
ANKS1A	chr6	84.05
PTPN23	chr3	84.05
SLC39A13	chr11	84.05
TAF11	chr6	84.05
AFTPH	chr2	83.13
AKIRIN2	chr6	83.13
RNF216	chr7	83.13
ZFPL1	chr11	82.44
BACH2	chr6	81.47
GUK1	chr1	81.12
ZNF212	chr7	80.92
TNFAIP3	chr6	80.82
YPEL3	chr16	80.51
RHBDD1	chr2	80.3
DPH1	chr17	80.26
KIAA1267	chr17	79.78
ZNF498	chr7	79.74
FEM1B	chr15	79.66
DNAJC9	chr10	79.37
BCL7A	chr12	79.35
ST6GALNAC2	chr17	79.05
APOL2	chr22	79.03

Gene Name	Chromosome	Fold Enrichment
C18orf8	chr18	79.03
EPM2A	chr6	79.03
G6PD	chrX	79.03
IKBKG	chrX	79.03
MINK1	chr17	79.03
PLA2G4C	chr19	79.03
EGR3	chr8	79.02
PGPEP1	chr19	79.02
MMAA	chr4	78.97
CD274	chr9	78.8
SENP8	chr15	78.8
POLH	chr6	78.51
QSOX1	chr1	78.51
LPGAT1	chr1	77.59
NFX1	chr9	77.59
PRKD2	chr19	77.59
PTPN9	chr15	77.59
SQSTM1	chr5	77.59
TBC1D22A	chr22	77.59
NUP214	chr9	77.39
RAD23A	chr19	76.57
ZBTB26	chr9	76.2
ZFAND2A	chr7	74.82
C15orf40	chr15	74.35
MAP3K5	chr6	73.86
PLEKHG1	chr6	73.86
SNORA43	chr9	73.32
C19orf12	chr19	72.74
CXXC4	chr4	72.74
E2F6	chr2	72.74
KIAA2018	chr3	72.74
KIAA2018	chr3	72.74
PEX14	chr1	72.74
PPP1R3D	chr20	72.74
PURA	chr5	72.74
SMYD4	chr17	72.35
TSNARE1	chr8	71.84
KIAA1267	chr17	71.46
IKBKG	chrX	71.12
KCNQ4	chr1	71.12
MTCP1	chrX	71.12
TMEM171	chr5	71.12
TOX4	chr14	71.12
ZBTB5	chr9	71.12
ZNF169	chr9	71.12
ZNF490	chr19	71.12
STK10	chr5	70.87
HIAT1	chr1	70.8
NAT15	chr16	70.53
SAMD4B	chr19	70.53
ZER1	chr9	70.31
AGBL2	chr11	70.25
CALCOCO2	chr17	70.25
DUSP16	chr12	70.25
KHDC1	chr6	70.25
KIF1C	chr17	70.25
MTMR1	chrX	70.25

Gene Name	Chromosome	Fold Enrichment
MUCL1	chr12	70.25
P2RX2	chr12	70.25
RNASET2	chr6	70.25
SAT1	chrX	70.25
ATP2A3	chr17	69.83
NAB2	chr12	69.83
KIAA1908	chr7	69.75
SNORA28	chr14	69.35
PRICKLE1	chr12	69.27
SON	chr21	69.27
DIP2B	chr12	69.26
CIC	chr19	68.46
ABCC3	chr17	67.89
ACP2	chr11	67.89
CHAF1A	chr19	67.89
GOSR2	chr17	67.89
RP9P	chr7	67.89
SLC25A36	chr3	67.89
TFE3	chrX	67.89
ZNF579	chr19	67.85
MCM4	chr8	67.14
RUNX1	chr21	67.14
XRN1	chr3	67.14
EP400	chr12	66.8
FAM82A2	chr15	66.36
TMSB10	chr2	66.03
ZMYM2	chr13	65.95
WIPI1	chr17	65.81
UNK	chr17	65.46
ABHD12	chr20	64.66
ANG	chr14	64.66
C6orf62	chr6	64.66
CCNT1	chr12	64.66
COMMD10	chr5	64.66
KLF4	chr9	64.66
MCOLN1	chr19	64.66
RUNX1T1	chr8	64.66
UBE2O	chr17	64.66
ZNF641	chr12	64.66
ATF3	chr1	64.43
VPS13C	chr15	64.41
SPTBN4	chr19	64.3
C20orf12	chr20	64.24
CECR6	chr22	64.01
C20orf29	chr20	63.73
FLYWCH1	chr16	63.41
IGFBP3	chr7	63.25
LOC643008	chr17	63.25
CELSR3	chr3	63.04
SPATA2	chr20	63.04
NFKBIA	chr14	62.63
CHP	chr15	62.52
PRKRIP1	chr7	62.35
GALNT10	chr5	62.07
MDM2	chr12	62.07
MED12L	chr3	62.07
WARS	chr14	62.07

Gene Name	Chromosome	Fold Enrichment
AKAP13	chr15	61.46
ANGEL2	chr1	61.46
BCAR3	chr1	61.46
C19orf28	chr19	61.46
C1QTNF6	chr22	61.46
C21orf91	chr21	61.46
C22orf32	chr22	61.46
C3orf38	chr3	61.46
DUSP8	chr11	61.46
ERCC5	chr13	61.46
IL15RA	chr10	61.46
LOC646762	chr7	61.46
PEX5L	chr3	61.46
PID1	chr2	61.46
QKI	chr6	61.46
RFFL	chr17	61.46
STAT5B	chr17	61.46
STIM2	chr4	61.46
TOX4	chr14	61.46
WTAP	chr6	61.46
WTAP	chr6	61.46
ZRANB2	chr1	61.46
DKFZp761E198	chr11	61.42
SLC34A3	chr9	61.42
MAP3K5	chr6	61.25
SPTBN4	chr19	60.96
FOXO3	chr6	60.91
SLC33A1	chr3	60.88
CREBZF	chr11	60.61
KAZALD1	chr10	60.61
MYLIP	chr6	60.52
PQLC2	chr1	60.2
STRADA	chr17	60.13
RAP2B	chr3	59.68
SAMD4A	chr14	59.68
SETD5	chr3	59.51
HIST3H2A	chr1	59.03
TM2D3	chr15	59.03
MT2A	chr16	59.02
MYO15B	chr17	58.88
CASP9	chr1	58.19
GCLC	chr6	58.19
GRB10	chr7	58.19
GUSB	chr7	58.19
HIC1	chr17	58.19
HIST1H2BD	chr6	58.19
LRRC1	chr6	58.19
NAB2	chr12	58.19
NDUFA11	chr19	58.19
NF1	chr17	58.19
PION	chr7	58.19
PPP3CC	chr8	58.19
PVRL1	chr11	58.19
SIK2	chr11	58.19
SLC20A1	chr2	58.19
SLC34A3	chr9	58.19
SNAP29	chr22	58.19

Gene Name	Chromosome	Fold Enrichment
ZNF695	chr1	58.19
ZNF70	chr22	58.19
AFF4	chr5	57.8
RAB1A	chr2	57.62
PAPOLA	chr14	57.49
DENND5B	chr12	57.47
UBE2O	chr17	57.47
NR3C1	chr5	57.31
TBL1X	chrX	57.16
LOC390940	chr19	57.05
N4BP1	chr16	57.05
XRCC1	chr19	57.05
CYTH1	chr17	56.57
HIST1H2AE	chr6	56.57
HIST1H2BG	chr6	56.57
PPAPDC1B	chr8	56.57
LRIG2	chr1	56.44
KIAA1530	chr4	56.33
TRRAP	chr7	56.31
TMCC1	chr3	56.25
DAPK2	chr15	56.22
OSBPL2	chr20	56.15
ZNF12	chr7	56.15
EDF1	chr9	56.1
SQSTM1	chr5	56.03
SAMD4A	chr14	55.95
TMEM214	chr2	55.89
FZR1	chr19	55.42
GLS	chr2	55.42
CD55	chr1	55.28
SGSM2	chr17	55.26
GCLC	chr6	54.96
XRCC5	chr2	54.96
DGKH	chr13	54.82
CELSR3	chr3	54.62
OTUD1	chr10	54.55
TADA2B	chr4	54.55
TMEM51	chr1	54.55
EPB41L4A	chr5	54.31
FLJ11235	chr5	54.31
IKZF2	chr2	54.31
LOC283922	chr16	54.31
PRSS23	chr11	54.31
RUNX2	chr6	54.31
SYNM	chr15	54.31
SIK2	chr11	54.16
PARP10	chr8	53.91
GPRC5C	chr17	53.88
RDH10	chr8	53.88
UNC93B1	chr11	53.88
PLXNC1	chr12	53.65
BDKRB2	chr14	53.52
CCDC75	chr2	53.34
FBXO34	chr14	53.34
IRF1	chr5	53.34
ZNF567	chr19	53.34
ZNF805	chr19	53.34

Gene Name	Chromosome	Fold Enrichment
APBB2	chr4	53.23
ACVR1	chr2	53.18
GNA13	chr17	53.18
PNKD	chr2	53.18
IRS2	chr13	53.15
RER1	chr1	53.13
CYLD	chr16	52.9
GPRC5C	chr17	52.9
HGS	chr17	52.9
IFT140	chr16	52.9
BATF3	chr1	52.68
CLCN6	chr1	52.68
FXR2	chr17	52.68
HIST1H1C	chr6	52.68
HS2ST1	chr1	52.68
HSD17B4	chr5	52.68
KCNMB3	chr3	52.68
LDOC1L	chr22	52.68
MEX3C	chr18	52.68
MTHFR	chr1	52.68
RGS2	chr1	52.68
SLC1A1	chr9	52.68
SLC35D2	chr9	52.68
TBC1D1	chr4	52.68
ZNF493	chr19	52.68
RAB8B	chr15	52.65
LSM14B	chr20	52.59
PSMB1	chr6	52.53
TINF2	chr14	52.53
ICA1	chr7	52.45
FOSL2	chr2	52.42
TOE1	chr1	52.37
NFE2L1	chr17	52.33
PSMB1	chr6	52.24
GTPBP2	chr6	52.22
LRRFIP1	chr2	52.22
MAD2L1BP	chr6	52.22
TLE1	chr9	52.22
CPLX2	chr5	52.04
LMTK3	chr19	52.04
PARP10	chr8	52.04
KCTD15	chr19	52.02
EPS8L1	chr19	51.98
ANG	chr14	51.72
FLJ90757	chr17	51.72
GIGYF2	chr2	51.72
IL15	chr4	51.72
MAP2	chr2	51.72
MAST4	chr5	51.72
PCCA	chr13	51.72
PTPN9	chr15	51.72
RNF38	chr9	51.72
SPATA2	chr20	51.72
TRIP11	chr14	51.72
WDR81	chr17	51.72
ZNF101	chr19	51.72
IKZF2	chr2	51.34

Gene Name	Chromosome	Fold Enrichment
PCGF2	chr17	51.34
STK40	chr1	51.34
ANGPTL6	chr19	51.28
CALCOCO1	chr12	51.21
PHF15	chr5	51.21
CIC	chr19	51.19
CES2	chr16	51.04
RPL32	chr3	51.04
C5orf45	chr5	50.98
CCNL1	chr3	50.92
FAM46A	chr6	50.8
KIAA0513	chr16	50.8
MAD2L1BP	chr6	50.8
STAT6	chr12	50.8
PWWP2B	chr10	50.78
P2RX4	chr12	50.6
TAP2	chr6	50.6
CDKN1A	chr6	50.43
EML2	chr19	50.43
GOLGA3	chr12	50.43
HERPUD1	chr16	50.43
MT1X	chr16	50.43
P2RY6	chr11	50.43
HIC1	chr17	50.28
RMND5B	chr5	50.1
UBC	chr12	50.1
SSH1	chr12	50.06
BATF	chr14	49.88
GMEB2	chr20	49.88
PTPN21	chr14	49.88
C15orf57	chr15	49.75
RHOQ	chr2	49.73
PIK3IP1	chr22	49.71
RBM12B	chr8	49.57
MEF2D	chr1	49.37
EDC4	chr16	49.22
DPH1	chr17	48.7
AKAP5	chr14	48.49
CTAGE5	chr14	48.49
DKFZp761E198	chr11	48.49
ELMO2	chr20	48.49
ERAP1	chr5	48.49
EXOC1	chr4	48.49
FANCC	chr9	48.49
HARBI1	chr11	48.49
ITGB4	chr17	48.49
LIPA	chr10	48.49
MCM4	chr8	48.49
MED12L	chr3	48.49
SLC25A28	chr10	48.49
SLC2A13	chr12	48.49
TACC1	chr8	48.49
TBL1X	chrX	48.49
TRIM26	chr6	48.49
UBC	chr12	48.49
ZNF41	chrX	48.49
TMC7	chr16	48.34

Gene Name	Chromosome	Fold Enrichment
FBXO32	chr8	48.03
GLTSCR2	chr19	47.26
MAN1A2	chr1	47.11
FLT3LG	chr19	47.02
MALAT1	chr11	46.95
CSAD	chr12	46.93
PSMB9	chr6	46.93
TAP1	chr6	46.93
FYN	chr6	46.71
EML2	chr19	46.55
FLJ42709	chr5	46.55
MAFK	chr7	46.55
NR2F1	chr5	46.55
SLC35E1	chr19	46.55
ZSWIM6	chr5	46.38
ANKHD1	chr5	46.27
CXCL16	chr17	46.24
ZMYND15	chr17	46.24
CECR6	chr22	46.18
SMAD3	chr15	46.18
ZFAND3	chr6	46.07
ACCS	chr11	45.93
ZNF34	chr8	45.85
AIDA	chr1	45.64
IKBKG	chrX	45.64
ITGA2	chr5	45.64
MAP3K12	chr12	45.64
ZZEF1	chr17	45.45
ZSWIM6	chr5	45.32
BRAP	chr12	45.29
C4orf33	chr4	45.26
CADM1	chr11	45.26
CD55	chr1	45.26
EVPL	chr17	45.26
FRAT2	chr10	45.26
G6PD	chrX	45.26
GLCCI1	chr7	45.26
IKBKG	chrX	45.26
LOC284454	chr19	45.26
OTUD5	chrX	45.26
PARP3	chr3	45.26
PPFIBP1	chr12	45.26
SFI1	chr22	45.26
SLC22A5	chr5	45.26
SRCAP	chr16	45.26
TFE3	chrX	45.26
TIMP2	chr17	45.26
TMTC3	chr12	45.26
XKR8	chr1	45.26
ZC3H12A	chr1	45.26
ZNF224	chr19	45.26
TAX1BP3	chr17	45.04
FOXO3	chr6	44.92
ANKS1A	chr6	44.76
TAF11	chr6	44.76
EGR4	chr2	44.57
DKFZp761E198	chr11	44.45

Gene Name	Chromosome	Fold Enrichment
EML2	chr19	44.45
NAB2	chr12	44.45
AAK1	chr2	44.34
EXT1	chr8	44.34
GALNT10	chr5	44.34
PTCH1	chr9	44.34
MED26	chr19	44.08
SLC2A13	chr12	44.08
ANKRD11	chr16	44.07
WDR61	chr15	44.03
DISC1	chr1	43.99
SRCAP	chr16	43.93
SLC48A1	chr12	43.92
C11orf63	chr11	43.9
CRLS1	chr20	43.9
GPR157	chr1	43.9
IP6K1	chr3	43.9
MTMR1	chrX	43.9
NPAT	chr11	43.9
SLFN5	chr17	43.9
ZNF396	chr18	43.9
CAMKK2	chr12	43.88
SLC26A6	chr3	43.8
TRIM26	chr6	43.76
CLPTM1	chr19	43.64
DUSP22	chr6	43.64
MAFG	chr17	43.64
MAST4	chr5	43.64
METTL14	chr4	43.64
NEURL3	chr2	43.64
REL	chr2	43.64
RERE	chr1	43.64
SPEF2	chr5	43.64
STX3	chr11	43.64
TMEM208	chr16	43.64
TSSC1	chr2	43.64
TTC15	chr2	43.64
UPF1	chr19	43.64
ZNF547	chr19	43.64
NADSYN1	chr11	43.54
C14orf132	chr14	43.48
GNAS	chr20	43.48
ABHD12	chr20	43.47
SPHK1	chr17	43.39
SH2D4B	chr10	43.25
RNF114	chr20	43.1
SKIL	chr3	43.1
SRCAP	chr16	43.1
TGFA	chr2	43.1
WDR8	chr1	43.1
TYSND1	chr10	42.94
FLJ32063	chr2	42.71
ABCD4	chr14	42.67
FRY	chr13	42.67
GIT2	chr12	42.67
MXD1	chr2	42.67
RNF103	chr2	42.67

Gene Name	Chromosome	Fold Enrichment
S100P	chr4	42.67
SLC36A1	chr5	42.67
PLEKHG2	chr19	42.58
IFI27L1	chr14	42.49
FOXN2	chr2	42.46
ELAVL2	chr9	42.43
FBXO34	chr14	42.43
SPINT1	chr15	42.43
XKR8	chr1	42.43
AARS2	chr6	42.32
PSME4	chr2	42.3
SMAD3	chr15	42.17
TCF7L1	chr2	42.17
VEPH1	chr3	42.17
ARFGAP1	chr20	42.03
BIN3	chr8	42.03
CHTF8	chr16	42.03
ERP44	chr9	42.03
FBF1	chr17	42.03
PHF20L1	chr8	42.03
ZNF34	chr8	42.03
HARS	chr5	41.95
HARS2	chr5	41.95
MTHFD2L	chr4	41.9
SSBP1	chr7	41.82
MAPK1IP1L	chr14	41.78
FAM48A	chr13	41.56
PVRL4	chr1	41.56
RBM39	chr20	41.56
ZNF689	chr16	41.56
IL15	chr4	41.51
LRIG2	chr1	41.43
C6orf62	chr6	41.31
C1orf104	chr1	41.22
CIAO1	chr2	41.22
VGLL4	chr3	41.22
RDH10	chr8	41.18
SRRM2	chr16	41.09
TNFSF13	chr17	41.03
ZNF704	chr8	41.03
ACHE	chr7	40.95
TBX3	chr12	40.95
C6orf1	chr6	40.89
KIAA0040	chr1	40.86
KIAA1967	chr8	40.84
MAFG	chr17	40.84
SPHK1	chr17	40.84
TAX1BP3	chr17	40.84
BACE2	chr21	40.67
EIF3B	chr7	40.67
HSD11B2	chr16	40.67
FLVCR2	chr14	40.6
C15orf48	chr15	40.5
TAF1C	chr16	40.48
FLJ90757	chr17	40.41
RPAP1	chr15	40.41
MST1R	chr3	40.13

Gene Name	Chromosome	Fold Enrichment
ZKSCAN5	chr7	40.03
ZNF394	chr7	40.03
CPSF4	chr7	40.01
CAST	chr5	39.93
IL4R	chr16	39.93
RFX1	chr19	39.93
TMUB2	chr17	39.8
CD46	chr1	39.65
MACC1	chr7	39.61
NDUFA9	chr12	39.57
CLDN1	chr3	39.26
ENTPD6	chr20	39.24
IDH1	chr2	39.1
SDC4	chr20	39.08
PASK	chr2	38.99
ABCD1	chrX	38.79
AFAP1	chr4	38.79
APBB2	chr4	38.79
BBX	chr3	38.79
BCL2L11	chr2	38.79
CA11	chr19	38.79
CCND3	chr6	38.79
CDC42EP3	chr2	38.79
DNAJC5B	chr8	38.79
DOK4	chr16	38.79
DUSP22	chr6	38.79
EXOC3	chr5	38.79
FAM46A	chr6	38.79
FXR2	chr17	38.79
HDAC5	chr17	38.79
HSD11B2	chr16	38.79
IFT81	chr12	38.79
IKBKG	chrX	38.79
IL15RA	chr10	38.79
IRF5	chr7	38.79
KIAA1217	chr10	38.79
MADD	chr11	38.79
METTL6	chr3	38.79
MKLN1	chr7	38.79
MNT	chr17	38.79
MNT	chr17	38.79
MX2	chr21	38.79
NMT1	chr17	38.79
OTUD5	chrX	38.79
PLXNA2	chr1	38.79
PPP1R3D	chr20	38.79
QSOX1	chr1	38.79
RAPGEF5	chr7	38.79
RBM10	chrX	38.79
REXO1	chr19	38.79
SCRN2	chr17	38.79
SEC24A	chr5	38.79
SERPINB1	chr6	38.79
SLC12A6	chr15	38.79
SLC35C2	chr20	38.79
SLK	chr10	38.79
STX3	chr11	38.79

Gene Name	Chromosome	Fold Enrichment
TAF8	chr6	38.79
TUG1	chr22	38.79
UBE2Z	chr17	38.79
WARS	chr14	38.79
XAB2	chr19	38.79
ZNF783	chr7	38.79
ZNFX1	chr20	38.79
FRAT2	chr10	38.75
H2AFJ	chr12	38.64
MED15	chr22	38.64
C8orf59	chr8	38.48
SPRY2	chr13	38.48
C4orf34	chr4	38.41
H2AFJ	chr12	38.41
AGXT2L2	chr5	38.21
SETX	chr9	38.11
LMF1	chr16	38.1
RAB24	chr5	38.1
WDFY2	chr13	37.97
ARL8A	chr1	37.95
FLJ44606	chr5	37.95
PDE4DIP	chr1	37.95
SCPEP1	chr17	37.95
RFFL	chr17	37.88
EHD1	chr11	37.82
LRRC8A	chr9	37.72
MOBKL2C	chr1	37.72
OGFR	chr20	37.72
STAP2	chr19	37.72
KIAA0226	chr3	37.62
RAB11FIP3	chr16	37.61
PSMC2	chr7	37.35
CIC	chr19	37.3
GIT2	chr12	37.3
INHBB	chr2	37.3
MYCBP2	chr13	37.3
RAP2B	chr3	37.3
RAB35	chr12	37.28
SQSTM1	chr5	37.14
LDLRAP1	chr1	37.07
EHD4	chr15	36.89
NIPBL	chr5	36.89
SMURF1	chr7	36.85
RAD9A	chr11	36.79
VEZF1	chr17	36.79
ZC3H3	chr8	36.75
PLEKHG2	chr19	36.7
LMAN2	chr5	36.64
PLCG2	chr16	36.6
AP4B1	chr1	36.58
DTNA	chr18	36.38
AMPH	chr7	36.37
C4orf10	chr4	36.37
CAV1	chr7	36.37
CEBPD	chr8	36.37
CEBPD	chr8	36.37
COL5A1	chr9	36.37

Gene Name	Chromosome	Fold Enrichment
MBOAT7	chr19	36.37
SPHK2	chr19	36.37
LMNA	chr1	36.27
ATXN7L1	chr7	36.21
ID3	chr1	36.12
EIF4E2	chr2	36.03
MTMR14	chr3	36.02
RTKN	chr2	36.02
C17orf53	chr17	35.98
DDR1	chr6	35.98
FAM160A1	chr4	35.92
FUT4	chr11	35.92
SIN3A	chr15	35.92
ARHGEF11	chr1	35.79
TMEM59	chr1	35.77
JAK2	chr9	35.73
MED12L	chr3	35.73
MYH14	chr19	35.73
FOXO3	chr6	35.7
C19orf21	chr19	35.64
MEF2D	chr1	35.63
C19orf50	chr19	35.59
GALNT10	chr5	35.56
GBA	chr1	35.56
MAF	chr16	35.56
NR4A3	chr9	35.56
PLD1	chr3	35.56
PRPF6	chr20	35.56
RBM38	chr20	35.56
TTC23	chr15	35.56
VRK3	chr19	35.56
TMEM67	chr8	35.52
CITED2	chr6	35.47
C20orf111	chr20	35.29
ASPHD2	chr22	35.27
PHF15	chr5	35.17
CARS	chr11	35.12
PPP2R2A	chr8	35.12
HBS1L	chr6	35.11
TRAM1	chr8	35.06
CEBPD	chr8	34.99
YPEL2	chr17	34.99
BBC3	chr19	34.91
GLI3	chr7	34.91
KCNAB2	chr1	34.91
LOC145783	chr15	34.91
MBD5	chr2	34.91
MKNK2	chr19	34.91
STIM2	chr4	34.91
TECPR2	chr14	34.91
UBN2	chr7	34.91
SIRT5	chr6	34.86
ZNF292	chr6	34.8
ANG	chr14	34.73
INPP4A	chr2	34.68
PTPN13	chr4	34.68
NR4A3	chr9	34.64

Gene Name	Chromosome	Fold Enrichment
RPL36	chr19	34.58
TRIM9	chr14	34.55
PDLIM7	chr5	34.48
TXNIP	chr1	34.48
DDIT3	chr12	34.41
LY6E	chr8	34.41
TSHZ3	chr19	34.23
LOC386597	chr2	34.14
BCL7A	chr12	33.94
C17orf48	chr17	33.94
CPNE8	chr12	33.94
ELAC1	chr18	33.94
EML4	chr2	33.94
FAM84A	chr2	33.94
KRBA1	chr7	33.94
MAP3K8	chr10	33.94
NEURL2	chr20	33.94
PARP3	chr3	33.94
PDS5B	chr13	33.94
RABL3	chr3	33.94
RGS9	chr17	33.94
SMARCA5	chr4	33.94
TPP1	chr11	33.94
TRAPPC10	chr21	33.94
TTC32	chr2	33.94
ZNF134	chr19	33.94
C14orf101	chr14	33.79
MTF2	chr1	33.69
TSSC1	chr2	33.67
TTC15	chr2	33.67
ANKRD10	chr13	33.6
FBXO7	chr22	33.59
TCF4	chr18	33.52
TMC8	chr17	33.44
ZNF444	chr19	33.44
C10orf76	chr10	33.36
TMC8	chr17	33.36
MLL4	chr19	33.34
SIX1	chr14	33.34
C7orf47	chr7	33.33
CLDN3	chr7	33.25
EGFR	chr7	33.25
KIAA1147	chr7	33.25
RGS16	chr1	33.25
SGCE	chr7	33.25
STAMBP	chr2	33.25
RHOC	chr1	33.12
LRRC50	chr16	32.97
SMPD3	chr16	32.97
ZFHX3	chr16	32.94
EHBP1L1	chr11	32.78
XPO1	chr2	32.68
TCP11L2	chr12	32.59
FAM19A5	chr22	32.35
ABCB6	chr2	32.33
APCDD1	chr18	32.33
BTK	chrX	32.33
C19orf12	chr19	32.33

Gene Name	Chromosome	Fold Enrichment
CD47	chr3	32.33
CYB561D1	chr1	32.33
DACT1	chr14	32.33
FANCC	chr9	32.33
FLOT1	chr6	32.33
GALNT10	chr5	32.33
HIF1AN	chr10	32.33
KCNIP2	chr10	32.33
LRRC8A	chr9	32.33
MAFG	chr17	32.33
MREG	chr2	32.33
NEURL3	chr2	32.33
OSMR	chr5	32.33
PDE5A	chr4	32.33
RNF4	chr4	32.33
SQRDL	chr15	32.33
TANC1	chr2	32.33
TNKS1BP1	chr11	32.33
TRIM4	chr7	32.33
VRK3	chr19	32.33
WDR55	chr5	32.33
WDR81	chr17	32.33
ZNF295	chr21	32.33
TNIP1	chr5	32.18
SMCR8	chr17	32.17
MAP3K7	chr6	32.12
RPS6KC1	chr1	32.12
CITED2	chr6	32.02
GTPBP3	chr19	31.98
MAFG	chr17	31.98
SGSH	chr17	31.86
GLTSCR2	chr19	31.81
ANKHD1	chr5	31.75
GNPTAB	chr12	31.74
KCNV1	chr8	31.74
LMTK3	chr19	31.58
WDR59	chr16	31.58
SKIL	chr3	31.52
ILF3	chr19	31.45
LOC284751	chr20	31.45
ITGA3	chr17	31.38
NADK	chr1	31.34
SNX5	chr20	31.34
ZNF101	chr19	31.34
MTMR4	chr17	31.33
AXIN2	chr17	31.28
MAP2K4	chr17	31.22
RUNX1	chr21	31.16
KCTD15	chr19	31.1
BOD1L	chr4	31.03
GMPPB	chr3	31.03
HOXD8	chr2	31.03
KIAA0754	chr1	31.03
RNF103	chr2	31.03
SLC46A3	chr13	31.03
TBK1	chr12	31.03
TTPAL	chr20	31.03
WSB1	chr17	31.03

Gene Name	Chromosome	Fold Enrichment
ZNF618	chr9	31.03
ZFAND3	chr6	30.94
ZKSCAN5	chr7	30.92
ZNF394	chr7	30.92
C6orf120	chr6	30.86
NR3C1	chr5	30.86
TM9SF4	chr20	30.76
RORC	chr1	30.71
MPRIP	chr17	30.63
SCYL1	chr11	30.63
LYPD3	chr19	30.48
PCGF5	chr10	30.48
UBN1	chr16	30.48
CALML4	chr15	30.41
ANGEL2	chr1	30.38
BCL9L	chr11	30.31
HRK	chr12	30.31
UBN1	chr16	30.31
SPTBN4	chr19	30.24
MON1B	chr16	30.17
ZFYVE1	chr14	30.17
ADAM17	chr2	29.98
SRPK2	chr7	29.95
PCSK7	chr11	29.94
PTGER3	chr1	29.84
TGIF1	chr18	29.84
C7orf47	chr7	29.69
RUNX2	chr6	29.52
FBXO7	chr22	29.46
NMNAT1	chr1	29.23
CUL7	chr6	29.21
PALB2	chr16	29.21
ARHGEF7	chr13	29.09
ATP6V1C2	chr2	29.09
C19orf52	chr19	29.09
CD55	chr1	29.09
CEACAM5	chr19	29.09
CLCN3	chr4	29.09
ELAVL1	chr19	29.09
GIGYF1	chr7	29.09
ICAM1	chr19	29.09
RHOQ	chr2	29.09
RIPK4	chr21	29.09
ZC3HAV1	chr7	29.09
ZNF333	chr19	29.09
ZNF766	chr19	29.09
MSL3	chrX	29.02
AKIRIN1	chr1	28.99
GTPBP2	chr6	28.89
C19orf50	chr19	28.86
LOC96610	chr22	28.86
EEF1D	chr8	28.83
GLCCI1	chr7	28.78
TMEM51	chr1	28.75
FAM110C	chr2	28.74
ETS2	chr21	28.61
TCF7L1	chr2	28.6

Gene Name	Chromosome	Fold Enrichment
PLXNB2	chr22	28.52
NADSYN1	chr11	28.24
ZNF226	chr19	28.22
ZHX2	chr8	28.21
STAT2	chr12	28.16
PLEKHG2	chr19	28.07
TTC32	chr2	28.05
CPM	chr12	28.02
DNAJC27	chr2	28.02
HSP90B1	chr12	28.02
MBD5	chr2	28.02
MICALL2	chr7	28.02
COG3	chr13	27.99
NGDN	chr14	27.97
SAMD4A	chr14	27.93
ZNF678	chr1	27.83
BLOC1S3	chr19	27.71
KCTD10	chr12	27.71
PGD	chr1	27.71
RANBP3	chr19	27.71
TRAPPC6A	chr19	27.71
UBE3B	chr12	27.71
ZFYVE19	chr15	27.71
ZCCHC6	chr9	27.56
SAMHD1	chr20	27.49
CD58	chr1	27.35
RHOC	chr1	27.35
UNK	chr17	27.35
ZNF223	chr19	27.35
APCDD1	chr18	27.16
CCRN4L	chr4	27.16
CLIP1	chr12	27.16
DNMBP	chr10	27.16
ELK1	chrX	27.16
GABARAPL3	chr15	27.16
HMOX1	chr22	27.16
ID3	chr1	27.16
JHDM1D	chr7	27.16
KIAA0556	chr16	27.16
LEPREL1	chr3	27.16
NDE1	chr16	27.16
PTPRK	chr6	27.16
RP9P	chr7	27.16
RPL32	chr3	27.16
TMEM165	chr4	27.16
WDR78	chr1	27.16
KLHL20	chr1	27.13
PDE4DIP	chr1	27.04
RARG	chr12	26.94
PVRL2	chr19	26.92
IL17RC	chr3	26.83
ZNF707	chr8	26.71
GABBR1	chr6	26.69
PXK	chr3	26.69
PDLIM7	chr5	26.67
RPUSD3	chr3	26.67
TEP1	chr14	26.67

Gene Name	Chromosome	Fold Enrichment
YPEL2	chr17	26.67
IGFBP3	chr7	26.62
GDF15	chr19	26.46
BRE	chr2	26.45
C2CD3	chr11	26.45
DLGAP4	chr20	26.45
SETD1A	chr16	26.45
ZNF277	chr7	26.45
SLC34A3	chr9	26.21
CBR3	chr21	26.19
ZBTB5	chr9	26.15
E2F6	chr2	26.11
SCUBE3	chr6	26.11
UFSP1	chr7	26.1
FAM53B	chr10	26.07
H2AFY	chr5	26.07
PDIA6	chr2	26
CLTC	chr17	25.99
UNK	chr17	25.93
TMEM165	chr4	25.89
ASB3	chr2	25.86
CBFA2T2	chr20	25.86
CLCF1	chr11	25.86
DOM3Z	chr6	25.86
FKBP15	chr9	25.86
IL6ST	chr5	25.86
LRRC50	chr16	25.86
OSBP	chr11	25.86
PASK	chr2	25.86
PIK3R1	chr5	25.86
PODXL	chr7	25.86
SPAG9	chr17	25.86
SPG20	chr13	25.86
STK19	chr6	25.86
TM9SF1	chr14	25.86
HSPB9	chr17	25.84
HNRNPA0	chr5	25.72
OGFOD2	chr12	25.72
HIST2H2BE	chr1	25.66
DLG4	chr17	25.62
TMEM144	chr4	25.47
TRIP10	chr19	25.44
FLJ10038	chr15	25.3
OTUB1	chr11	25.29
HSP90B1	chr12	25.22
MAF	chr16	25.22
MAX	chr14	25.22
CMYA5	chr5	25.19
VTI1A	chr10	25.14
FDXR	chr17	25.03
ORAI2	chr7	25.02
MTMR4	chr17	25
EEF1D	chr8	24.94
GDPD1	chr17	24.94
GIGYF1	chr7	24.94
HOXB9	chr17	24.94
MEF2C	chr5	24.94

Gene Name	Chromosome	Fold Enrichment
STAP2	chr19	24.94
TMEM219	chr16	24.94
ARPC5	chr1	24.88
FOSL2	chr2	24.87
SAP30L	chr5	24.79
ATP6V1G1	chr9	24.74
C20orf111	chr20	24.71
PVRL2	chr19	24.47
RBM7	chr11	24.3
B4GALNT1	chr12	24.25
CAMTA2	chr17	24.25
ITGA3	chr17	24.25
PAQR7	chr1	24.25
PELI1	chr2	24.25
SEPSECS	chr4	24.25
SIN3A	chr15	24.25
TMEM63A	chr1	24.21
TREX1	chr3	24.17
ATXN7	chr3	24.16
TIA1	chr2	24
STX12	chr1	23.97
RABAC1	chr19	23.95
XPO1	chr2	23.87
TMEM40	chr3	23.83
SMAD9	chr13	23.71
ZNF629	chr16	23.71
STK40	chr1	23.68
RFK	chr9	23.66
TNFRSF10B	chr8	23.49
POP7	chr7	23.44
ATG2B	chr14	23.43
EHBP1L1	chr11	23.41
GNAS	chr20	23.41
ALCAM	chr3	23.28
CXCR7	chr2	23.28
EGLN2	chr19	23.28
KIAA0226	chr3	23.28
MCL1	chr1	23.28
MGAT4A	chr2	23.28
MIER3	chr5	23.28
RNF149	chr2	23.28
TSNARE1	chr8	23.28
ZHX2	chr8	23.28
HEXB	chr5	22.94
NCF4	chr22	22.87
INHBB	chr2	22.82
TGIF1	chr18	22.82
ABCC5	chr3	22.71
RARG	chr12	22.7
RFX1	chr19	22.63
RORC	chr1	22.63
SHC1	chr1	22.63
ZNF564	chr19	22.63
THOC5	chr22	22.62
SGCE	chr7	22.6
FLOT1	chr6	22.52
ADHFE1	chr8	22.46

Gene Name	Chromosome	Fold Enrichment
C1orf54	chr1	22.38
ARNT	chr1	22.31
C6orf120	chr6	22.17
ITPKC	chr19	22.17
MLLT3	chr9	22.17
RAB4B	chr19	22.17
RHCG	chr15	22.17
STAP2	chr19	22.17
PCGF2	chr17	22.11
ABHD11	chr7	22.1
FLVCR2	chr14	22.07
CAMTA2	chr17	22.05
NLRC5	chr16	22.04
BCL2L13	chr22	21.96
INHBB	chr2	21.92
ABHD11	chr7	21.82
CHTF8	chr16	21.82
CPNE5	chr6	21.82
ATP6V0C	chr16	21.77
PSMG4	chr6	21.76
RARG	chr12	21.68
MVP	chr16	21.63
DERL1	chr8	21.55
DMWD	chr19	21.55
NPC1	chr18	21.55
RARG	chr12	21.55
SEZ6L2	chr16	21.55
UBR2	chr6	21.55
UBA6	chr4	21.53
NAMPT	chr7	21.44
EDC3	chr15	21.43
ZFYVE20	chr3	21.43
NQO1	chr16	21.41
ATP6V1H	chr8	21.4
INSM2	chr14	21.34
LGALS8	chr1	21.34
TUG1	chr22	21.31
STX4	chr16	21.29
CASP9	chr1	21.26
C16orf91	chr16	21.19
GEM	chr8	21.16
KCNV1	chr8	21.16
MAPK3	chr16	21.16
GRN	chr17	21.14
KCNG1	chr20	21.13
LSM14B	chr20	21.08
C14orf101	chr14	21.01
ITPRIPL2	chr16	21.01
GIT2	chr12	20.95
DLX3	chr17	20.82
MRPL10	chr17	20.78
EIF2S2	chr20	20.75
PMAIP1	chr18	20.72
NUPR1	chr16	20.68
SRRM2	chr16	20.66
GATA6	chr18	20.61
PPA1	chr10	20.6
ISG20	chr15	20.56

Gene Name	Chromosome	Fold Enrichment
PDE5A	chr4	20.54
RNF43	chr17	20.48
ALCAM	chr3	20.46
PARP6	chr15	20.39
PLCG1	chr20	20.35
MYO9B	chr19	20.33
PPL	chr16	20.3
PARD6B	chr20	20.11
BCL2L11	chr2	20.08
OGFOD2	chr12	20
TFPI2	chr7	19.97
HIST2H2BE	chr1	19.89
PNKD	chr2	19.79
TADA2B	chr4	19.55
Sep-04	chr17	19.4
AGXT2L2	chr5	19.4
ATOH8	chr2	19.4
ATP6V1B1	chr2	19.4
BAX	chr19	19.4
BLVRA	chr7	19.4
BTRC	chr10	19.4
C10orf28	chr10	19.4
CDC42EP3	chr2	19.4
CNPY2	chr12	19.4
FOSL2	chr2	19.4
FOSL2	chr2	19.4
GABBR1	chr6	19.4
GNA13	chr17	19.4
HYOU1	chr11	19.4
ITPRIPL2	chr16	19.4
KCNH8	chr3	19.4
MARS	chr12	19.4
NAV1	chr1	19.4
NOD1	chr7	19.4
ODZ4	chr11	19.4
PDIA4	chr7	19.4
RBM25	chr14	19.4
REXO1	chr19	19.4
SEZ6L2	chr16	19.4
ZNF804A	chr2	19.4
SHC1	chr1	19.38
DLX3	chr17	19.05
POLR3C	chr1	19
RNF115	chr1	19
ZFAND3	chr6	18.96
ANG	chr14	18.94
MTMR11	chr1	18.94
CIC	chr19	18.86
HERC5	chr4	18.85
SIN3A	chr15	18.65
ZNF564	chr19	18.62
PRKAR1A	chr17	18.49
SAR1A	chr10	18.38
ZFHX3	chr16	18.26
C1QL1	chr17	18.16
LOC643008	chr17	18.1
MUC1	chr1	18.04
ZNF408	chr11	18.01

Gene Name	Chromosome	Fold Enrichment
PSMC2	chr7	18
KCNH8	chr3	17.97
BTG1	chr12	17.9
MUC1	chr1	17.9
C5orf41	chr5	17.88
RBCK1	chr20	17.86
GABBR1	chr6	17.78
SOX13	chr1	17.69
DDIT3	chr12	17.65
HBP1	chr7	17.63
RABEP2	chr16	17.63
TNIK	chr3	17.62
CHMP4C	chr8	17.46
KIAA0226	chr3	17.46
LRRC1	chr6	17.46
RASGRF1	chr15	17.46
SLC37A1	chr21	17.46
TGIF1	chr18	17.46
TRIM56	chr7	17.46
MGEA5	chr10	17.38
BATF	chr14	17.32
C1orf21	chr1	17.24
RECQL5	chr17	17.24
SLC22A15	chr1	17.18
PGCP	chr8	17.11
PANK2	chr20	17.01
DFNB31	chr9	16.97
DISP1	chr1	16.97
RBM4	chr11	16.97
ZDHHC13	chr11	16.97
CELSR2	chr1	16.88
POP7	chr7	16.85
MVP	chr16	16.81
TTLL5	chr14	16.81
C13orf15	chr13	16.63
CCND3	chr6	16.63
PSMA3	chr14	16.63
PSME2	chr14	16.6
RNF31	chr14	16.6
WDR81	chr17	16.56
IL17RC	chr3	16.16
TAPBP	chr6	16.09
PSMB4	chr1	15.89
COPA	chr1	15.87
NCSTN	chr1	15.87
SETD1A	chr16	15.87
NFKBIB	chr19	15.59
ZNF260	chr19	15.59
FAM174B	chr15	15.54
RAD51L1	chr14	15.52
ZFYVE26	chr14	15.52
TGOLN2	chr2	15.48
MICALL2	chr7	15.31
ABHD11	chr7	15.26
ARAP1	chr11	15.09

Gene Name	Chromosome	Fold Enrichment
DGKH	chr13	14.55
TRPM7	chr15	14.55
TTLL6	chr17	14.55
PARD6B	chr20	14.52
FAM83A	chr8	14.39
TPCN2	chr11	14.31
WDR8	chr1	14.26
ZNF277	chr7	14.14
CHST11	chr12	14.11
TRIB1	chr8	14.08
ZDHHC7	chr16	13.55
SGCE	chr7	13.13
NOSIP	chr19	13.12
UBC	chr12	13.03
DLX3	chr17	12.49
DAPK2	chr15	12.47
MAPKAPK2	chr1	12.34
SLC7A8	chr14	12.34
ZMYND8	chr20	12.27
BCAS1	chr20	11.78
CYP19A1	chr15	11.75
DENND2C	chr1	11.74
ZMYND8	chr20	11.44
GAB1	chr4	10.51
ZMYND8	chr20	9.15
SNX11	chr17	8.62